KNOWLEDGE CONSTRUCTION IN HEALTH SUPPORT GROUP ONLINE DISCUSSIONS

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Management Information Systems

School of Management, Massey University, Palmerston North, New Zealand

Achmad Ghazali

2012

ABSTRACT

The ongoing transition to the patient-centred healthcare paradigm suggests that patients adopt an active role in managing their health conditions. As the result, the Internet is becoming an important source of health-related information. Internet-based health support groups allow patients to access diverse information relevant to their particular situation by participating in online discussions. The quality of such information may have effects on the patients' outcomes.

According to social constructivism, knowledge in online discussions is constructed in interactions between the individuals involved, as recommendations made over the discussion are clarified and scrutinized. Therefore, knowledge construction is likely to affect the quality of health-related information generated in health support group online discussions.

The purpose of the present study was to investigate the effects of knowledge construction in health support group online discussions on perceived information quality, information quality from the perspective of information consumers, and on information integrity, information validity from the point of view of the current state of scientific knowledge. It was hypothesized that knowledge construction results in better perceived information quality and in higher information integrity.

A health support group online discussion site devoted to weight management was used as a source of data. Quantitative content analysis was used, with a discussion thread as a unit of analysis.

Knowledge construction was operationalized as a two-dimensional construct with the dimensions of explicitation (lower level knowledge construction activities) and evaluation (higher level knowledge construction activities). The coding scheme was based on the prior studies of knowledge construction in the field of e-Learning. Perceived information quality was operationalized by adapting an existing measure from survey-based research. Information integrity was operationalized by using a simplified Delphi technique—health-related recommendations were extracted from the discussion content by coders and were assessed by domain experts.

Explicitation was found to affect perceived information quality with a medium effect size. Evaluation did not affect perceived information quality, and information integrity was not affected by any of the dimensions of knowledge construction.

Thus, low level knowledge construction contributed to perceived information quality, resulting in health-related information that is more relevant and useful from the perspective of its consumers. Nonetheless, knowledge construction activities were not found to result in higher prevalence of scientifically sound recommendations.

Based on the findings, the study suggested that moderators of health support group online discussions should promote explicitation by encouraging clarifications and refinements of health-related recommendations. Moreover, participation of qualified health practitioners is desirable to promote health-related behaviours based on evidence-based knowledge and to expose recommendations that have uncertain or even dangerous effects.

Acknowledgments

I thank my supervisors, Dr. Alexei Tretiakov, Dr. Inga Hunter, and Dr. Barbara Crump, who provided valuable insights, suggestions, and encouragements.

I thank the examiners for their comments and constructive critique, which helped me to better understand the implications of my study and to improve the thesis overall.

Two institutions, School of Business and Management at Bandung Institute of Technology (SBM-ITB) and Putera Sampoerna Foundation (PSF), supported my research. I thank SBM-ITB, where I am pursuing my teaching career, for giving me the support and opportunity to continue my study abroad. I am so grateful to PSF for providing me financial support throughout my doctoral studies. I would not have been able to complete my studies without the support of both institutions.

This acknowledgement would be incomplete without mentioning my family: my beloved wife Myrna, my beloved son Haris, my beloved daughter Hanna, my beloved parents, and my beloved parents-in-law for their continuous support, encouragement, and patience. To each one of you, I humbly say thank you.

Table of Contents

Acknowledgment iv

List of Figures xii

List of Tables xiv

List of Abbreviations xvi

Chapter 1. Introduction 1

1.1 Background of the Study 1

1.2 Purpose of the Study 3

1.3 Significance of the Study 3

1.3.1 Significance to Theory 4

1.3.2 Significance to Practice 4

1.4 Definition of Terms 5

1.5 Theoretical Framework 6

1.6 High-Level Hypotheses 7

1.7 Summary of Methods 10

1.8 Organization of the Thesis 10

Chapter 2. Literature Review 12

2.1 Introduction 12

2.2 Method Used to Conduct the Literature Review 12

2.3 Patient-Centered Care and Online Health Support Groups 15

2.4 Lurkers 18

2.5 Constructivism and Knowledge Construction 19

2.5.1 Individual Constructivism 19

2.5.2 Social Constructivism 19

2.6 Conceptualizations and Measures of Knowledge Construction in Online Discussions 20

2.6.1 Henri (1992)—Learning Process Framework 21

2.6.2 Gunawardena, Lowe, and Anderson (1997)—Knowledge Construction in CMC Environments 25

2.6.3 Veerman, Andriessen, and Kanselaar (1999)—Collaborative Learning and Argumentation 28

2.6.4 Garrison (1991)—Cycle of Learning 36

2.6.5 Other Studies 41

2.6.6 Summary and Comparison of the Measures of Knowledge Construction 43

2.7 Knowledge Construction in KM 50

2.7.1 McAdam and McCreedy Model of the Role of Knowledge Construction in KM 51

2.7.2 Enhancing Knowledge Construction at Organizations 52

2.7.3 An Empirical Study of Online Discussions Conducted From a KM Perspective 53

2.8 Perceived Information Quality and Information Integrity 54

2.9 Information Integrity and Evidence-Based Medicine 59

2.10 Knowledge Gap 62

2.11 Health Literacy 63

2.11.1 Conceptualization of Health Literacy 63

2.11.2 Health Support Online Discussions as an Antecedent of Health Literacy 64

2.11.3 Health Literacy as an Antecedent of Knowledge Construction 67

2.11.4 Health Literacy and the Purpose of the Present Study 67
2.12 Summary 69
Chapter 3. Model Development and Hypotheses 72
3.1 Introduction 72
3.2 Conceptualization of Constructs 72
3.2.1 Knowledge Construction 73
3.2.2 Perceived Information Quality 75
3.2.3 Information Integrity 78
3.3 Structural Model and Detailed Hypotheses 78
3.3.1 Knowledge Construction Results in Better Perceived Information Quality (H1) 79
3.3.2 Knowledge Construction Results in Better Information Integrity (H2) 81
3.4 Summary 82
Chapter 4. Research Methodology 83
4.1 Introduction 83
4.2 Overview of Research Procedures 84
4.3 Overall Approach to Research 84
4.3.1 Positivism Versus Interpretivism 84
4.3.2 Quantitative Research Versus Qualitative Research 87
4.3.3 Explanatory Research Versus Exploratory Research 87
4.3.4 Approach to Data Collection 88
4.4 Approach to Content Analysis 89
4.4.1 Reaffirming the Suitability of Content Analysis 89

4.4.2 Framework for Content Analysis 91
4.4.3 Data Making 91
4.4.4 Defining the Semantics of Data 92
4.4.5 Operationalization as Data Language 93
4.5 Choice of Content Source 94
4.6 Unitizing 95
4.7 Sampling and Screening 96
4.8 Coding 102
4.8.1 Overview of Coding Procedures 102
4.8.2 Coder Qualifications 103
4.8.3 Coding Scheme Development, Coder Training, and Coding Procedures 105
4.8.4 Operationalization of Knowledge Construction 107
4.8.5 Operationalization of the Dimensions of Perceived Information Quality 108
4.8.6 Operationalization of Information Integrity 109
4.8.7 Identification of Task Related Threads and Coding Procedures for Information Integrity 110
4.9 Overall Approach to Data Analysis 111
4.10 Reliability 112
4.10.1 Reliability of Coding 112
4.10.2 Reliability of Measures of Latent Constructs 113
4.11 Validity 113
4.11.1 Face Validity 113

4.11.2 Content Validity 114
4.11.3 Nomological Validity 114
4.11.4 Construct, Convergent, and Discriminant Validity 115
4.11.5 Internal Validity 115
4.11.6 External Validity (Generalizability) 116
4.11.7 Ecological Validity 116
4.11.8 Social Validity 117
4.12 Procedures Used to Assess Convergent and Discriminant Validity 117
4.13 Approach to Hypotheses Testing 118
4.14 Assumptions of Statistical Tests 119
4.15 Software Programs Used for Statistical Analysis 120
4.16 Summary 120
Chapter 5. Results and Analysis 123
5.1 Introduction 123
5.2 Descriptive Statistics 123
5.3 Inter-Rater Reliability 126
5.4 Normality 129
5.5 Exploratory Factor Analysis 131
5.6 PLS Measurement Model 138
5.7 PLS Structural Model 143
5.8 Post-Hoc Analyses 144
5.8.1 A Model With Both Knowledge Construction and Perceived Information Quality Treated as One-Dimensional Constructs 144

5.8.2 A Model With Explicitation and Evaluation as Separate Constructs and With Perceived Information Quality Treated as a One-Dimensional Construct 147

5.8.3 A Model With Explicitation Hypothesized to Affect Evaluation 149

5.8.4 A Model of Coder's Conceptions 152

5.8.5 Information Integrity: Analysis at the Level of a Recommendation 155

5.9 Interpretation of Data Analysis Results 157

5.9.1 Conceptualization of Constructs 157

5.9.2 Support of High-Level Hypotheses 160

5.10 Summary 163

Chapter 6. Discussion and Conclusion 166

6.1 Introduction 166

6.2 Overview of the Study 166

6.3 Implications of the Study 174

6.3.1 Implications for Theory 174

6.3.2 Implications for Practice 180

6.4 Limitations of the Study 190

6.5 Suggestions for Future Research 192

6.6 Conclusions 194

Appendix A. Knowledge Construction Form 197

Appendix B. Perceived Information Quality Form 200

Appendix C. Information Integrity Data Extraction Form 202

Appendix D. Information Integrity Expert Assessment 204

Appendix E. Details of Post-Hoc Analyses 219

E.1 A Model With Both Knowledge Construction and Perceived Information Quality Treated as One-Dimensional Constructs—The Details of Measurement Model Analysis 219

E.2 A Model With Explicitation and Evaluation as Separate Constructs and With Perceived Information Quality Treated as a One-Dimensional Construct—The Details of Measurement Model Analysis 222

E.3 A Model With Explicitation Hypothesized to Affect Evaluation—The Details of Measurement Model Analysis 225

E.4 A Model of Coder's Conceptions—The Details of Measurement Model Analysis 228

Appendix F. Mplus Code for the EFA Analysis Discussed in Section 5.5, With the Number of Factors Set to Four 238

Appendix G. Mplus Code for Linear Regression Analysis Discussed in Section 5.8.5 239

References 240

List of Figures

Figure 1. Knowledge construction in an online discussion. 7

Figure 2. High-level research model. 8

Figure 3. Mapping between Veerman and Veldhuis-Diermanse (2001) and Gunawardena et al. (1997) schemes. 33

Figure 4. Mapping between Veerman and Veldhuis-Diermanse (2001) and Fahy et al. (2000) schemes. 35

Figure 5. Cycle of learning model of knowledge construction by Garrison, based on Garrison (1991, p. 293). 37

Figure 6. Knowledge construction in knowledge management—A model by McAdam and McCreedy (1999) (p. 103). 52

Figure 7. The focus of the present study and health literacy. The constructs and the hypothesis falling within the scope of the present study are shown with bold lines. 68

Figure 8. Research model from section 1.6 (Figure 2) with knowledge construction represented by explicitation and evaluation, and perceived information quality—by relevance, understandability, and usefulness. 78

Figure 9. The main steps of this study. 84

Figure 10. Definition of "typical" threads (given by the shaded area). 97

Figure 11. Distribution of threads by the number of messages in a thread. Bins corresponding to typical threads are shown in grey. 98

Figure 12. Distribution of threads by the number of contributors. Bins corresponding to typical threads are shown in grey. 99

Figure 13. Distribution of task related threads by the number of messages in a thread. 101

Figure 14. Distribution of task related threads by the number of contributors. 101

Figure 15. Coding procedures. 103

Figure 16. Coding information integrity. 110

Figure 17. Scree test plot for the indicators of explicitation, evaluation, relevance, and usefulness. 131

Figure 18. Scree test plot for the indicators of the dimensions of knowledge construction: explicitation and evaluation. 135

Figure 19. Scree test plot for the indicators of the dimensions of perceived information quality: relevance and usefulness. 136

Figure 20. Structural model for PLS analysis. 138

Figure 21. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines. 144

Figure 22. Structural model with both knowledge construction and perceived information quality treated as one-dimensional constructs. 145

Figure 23. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines. 146

Figure 24. Structural model with perceived information quality treated as a one-dimensional construct. 148

Figure 25. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines. 149

Figure 26. Structural model with the hypothesis that explicitation affects evaluation. 150

Figure 27. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines. 151

Figure 28. Structural model accounting for coders' conceptions. 153

Figure 29. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. For paths between higher-level constructs and their coder's conceptions, all *p* values were below 0.001, and are not shown in the figure. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines. 154

Figure 30. Nomological framework around knowledge construction in online discussions. The part of the framework addressed by the present study is shown in a box with a dashed border. 176

List of Tables

Table 2-1 *Numbers of Citations of Articles Introducing Approaches to Conducting a Literature Review* 15

Table 2-2 *Henri's (1992) Dimensions of Learning* 21

Table 2-3 *Bullen's (1998) Scale for Critical Thinking* 22

Table 2-4 *Inter-Rater Reliability in Hara et al. (2000)* 23

Table 2-5 *Categories Retained by Pena-Shaff and Nicholls'(2004) That Matched Levels and Categories in the Henri's (1992) Model* 24

Table 2-6 *Gunawardena et al.'s Phases of Knowledge Construction* 26

Table 2-7 *Veerman et al.'s (1999) Dimensions* 28

Table 2-8 *Veerman et al.'s (1999) Inter-Rater Reliability* 29

Table 2-9 *Veerman and Veldhuis-Diermanse's (2001) Dimensions* 30

Table 2-10 *The Numbers of Discussions and Messages Used for Analysis in Each Study* 31

Table 2-11 *Garrison et al.'s (1999) Dimension of Cognitive Presence* 37

Table 2-12 *The Structure of the Newman, Johnson, Webb, and Cochrane (1997) Model of the Quality of Learning* 38

Table 2-13 *Transcripts Used in Garrison et al. (2001) Study* 39

Table 2-14 *Inter-Rater Reliability in Garrison et al. (2001) Study* 40

Table 2-15 *Knowledge Construction Scales Used by Schrire (2004)* 40

Table 2-16 *Coding Scheme by Järvelä and Häkkinen (2002)* 41

Table 2-17 *Coding Scheme for Shared Understanding (Puntambekar, 2006)* 43

Table 2-18 *Synthesis of Measures of Knowledge Construction* 46

Table 2-19 *Inter-Rater Reliability in Studies of Knowledge Construction* 48

Table 3-1 *The Content of the Explicitation Construct* 74

Table 3-2 *The Content of the Evaluation Construct* 75

Table 3-3 *Dimensions of Perceived Information Quality* 77

Table 4-1 *Approach to Content Analysis in the Present Study in Terms of Krippendorff's (2004) Framework for Content Analysis* 92

Table 4-2 *Screening of Threads* 100

Table 4-3 *Coders* 105

Table 4-4 *Operationalization of the Dimensions of Knowledge Construction* 107

Table 4-5 *Operationalization of the Dimensions of Perceived Information Quality* ... 108

Table 5-1 *Descriptive Statistics for Knowledge Construction* 124

Table 5-2 *Descriptive Statistics for Perceived Information Quality* 125

Table 5-3 *Descriptive Statistics for Information Integrity* 126

Table 5-4 *Inter-Rater Reliability for the Indicators of the Dimensions of Knowledge Construction* 127

Table 5-5 *Inter-Rater Reliability for the Indicators of the Dimensions of Perceived Information Quality* 128

Table 5-6 *Inter-Rater Reliability for Information Integrity* 129

Table 5-7 *Kurtosis and Skewness* 130

Table 5-8 *Factor Loadings in EFA for all Indicators of Explicitation, Evaluation, Relevance, and Usefulness, with the Number of Factors Set to Four* 132

Table 5-9 *Factor Loadings in EFA for all Indicators of Explicitation, Evaluation, Relevance, and Usefulness, with the Number of Factors Set to Two* 134

Table 5-10 *Factor Loadings in EFA for Indicators of Explicitation and Evaluation, with the Number of Factors Set to Two* 136

Table 5-11 *Factor Loadings in EFA for Indicators of Relevance and Usefulness, with the Number of Factors Set to Two* 137

Table 5-12 *Indicator Loadings on Dimensions of Knowledge Construction* 139

Table 5-13 *Indicator Loadings on Dimensions of Perceived Information Quality* 139

Table 5-14 *Indicator Loadings on Information Integrity* 140

Table 5-15 *Internal Consistency Indices* 140

Table 5-16 *Average Variance Extracted* 141

Table 5-17 *Latent Variables Correlation Matrix* 143

List of Abbreviations

AVE	Average Variance Extracted
CMC	Computer Mediated Communication
CR	Coefficient of Reliability
CSCL	Computer Supported Collaborated Learning
EFA	Exploratory Factor Analysis
KM	Knowledge Management
KMS	Knowledge Management System
PLS	Partial Least Squares
SD	Standard Deviation
SEM	Structural Equation Modelling

Chapter 1. Introduction

1.1 Background of the Study

According to a survey conducted in 2008 by the Pew Research Centre, 83% of internet users in the United States obtained information related to health and healthcare via online media (Fox & Jones, 2009). Furthermore, 60% of patients (people diagnosed with a health condition) who used Internet for seeking health information joined online health support groups to exchange health-related information with other patients. These numbers indicate that the Internet is a major source of health-related information, rather than merely a supplement.

Internet usage has enabled patients to access health-related information that was previously unreachable (Jadad & Gagliardi, 1998). For example, people can access health-related information using the Cochrane library, which is available over the Internet (http://www.thecochranelibrary.com). Yet, simply obtaining information by searching online databases does not meet all of the information needs of patients or other people with health-related concerns (Jadad & Gagliardi, 1998). Sometimes, they need to discuss their health-related problems with others, to obtain advice tailored to their particular situation or to gain moral support. Therefore, they join health support groups and participate in online discussions.

A community of practice is a group of people sharing an interest in a particular topic who participate in discussions intended to exchange and build up knowledge on the topic (Cox, 2005; Johnson, 2001; Wenger, 1998). People in a community of practice gain knowledge by sharing their ideas, experience, and information. Communities of practice have been using online discussions to communicate and to build knowledge for more than twenty years (Butler, Sproull, Kiesler, & Kraut, 2002; Maloney-Krichmar & Preece, 2005).

From the knowledge management (KM) point of view, online health support groups are communities of practice (Dalkir, 2011). This is because the members of an online health support group have similar interests (relating to a specific health issue), and

because they share knowledge and experiences with each other (Davidson & Voss, 2002).

Interactions among members of a health support group participating in online discussions are similar to interactions in online discussion groups set up for the purposes of e-Learning; discussion groups in e-Learning have been studied extensively (Puntambekar, 2006; Schellens & Valcke, 2005; Veerman & Veldhuis-Diermanse, 2001). The concept of knowledge construction stems from social constructivism theory, a theory that suggests that learning results from social interaction and that knowledge is co-constructed when individual learners negotiate meaning in interactions with each other and with experts (Hmelo-Silver, 2003; Leahey & Harris, 1985; Palincsar, 1998). In online discussions individuals gain better understanding by sharing their knowledge with others because the feedback from others enriches their knowledge.

KM views knowledge construction as a KM activity—the connectionist view of KM (which relies on social constructivist as one of its theoretical antecedents) suggests that knowledge construction results from interaction among employees at an organization (Dalkir, 2011). Knowledge development in KM is adding to knowledge available at the organization by making sense of it in view of the constantly changing local context. Knowledge development occurs via knowledge construction, with individual employees sharing their knowledge and discussing the knowledge shared, thus making it meaningful in view of the current situation. The resulting contextualized knowledge is internalized by the employees participating in discussions, resulting in the overall improvement of the organization's knowledge base.

From the social constructivism point of view, health support group online discussions result in learning via knowledge construction within the group; nonetheless, the nature of discussions in health support groups is different from that of online discussions in e-Learning (White & Dorman, 2001). Online discussions in health support groups are conducted informally (unlike in e-Learning, there is no externally imposed agenda), and contributors' backgrounds vary very broadly (there are no inclusion criteria, such as entrance exams in academia) resulting in a broad distribution in contribution quality. Moreover, sometimes an expert in the field of interest is not available (which may affect the information integrity of the content), and there are no external incentives to

contribute to the discussion (thus, it is particularly likely that many of the participants only read and never contribute). At the same time, the knowledge gained in online health support groups may have an immediate effect on the wellbeing of the individuals involved. Therefore, even though discussions in e-Learning and discussions in online health support groups are, in some respects, similar, the results of the numerous studies of knowledge construction in discussions in e-Learning cannot be immediately generalized to health support group online discussions, necessitating an empirical study of knowledge construction in health support group online discussions.

1.2 Purpose of the Study

In education, empirical studies of knowledge construction in online discussions are well established. These studies provide evidence of learning via knowledge construction in online discussions (see section 2.6 for a review of literature on knowledge construction in e-Learning discussions).

In knowledge management, even though the concept of knowledge construction via online discussions is well established, empirical studies of knowledge construction in online discussions in knowledge management are rare and limited (see section 2.7 for a review of literature on knowledge construction in knowledge management).

It is of interest to investigate knowledge construction in online discussions in a non-educational context. Health support group online discussions are particularly interesting because they are widespread (as suggested by the literature review presented in section 2.3) and, with the Facebook generation coming of age, likely to grow. In particular, in view of the potential effect of the quality of information gained in health support group online discussions on the wellbeing of individuals, it is of importance to investigate if knowledge construction contributes to the quality of information gained via such discussions.

1.3 Significance of the Study

This section explores the significance of this study from the perspectives of theory and practice.

1.3.1 Significance to Theory

Even though knowledge construction is a concept frequently referred to in the studies of online discussions, the studies involving measuring knowledge construction overwhelmingly focus on describing knowledge construction, rather than on testing the relationships of knowledge construction to other constructs. By testing the effects of knowledge construction on information quality, this study contributes to the nomological framework around knowledge construction, thus opening the way for further quantitative studies of the effects and antecedents of knowledge construction in online discussions.

Even though knowledge construction in online discussions is relevant to learning in general (such as learning as part of knowledge management at organizations), the existing studies of knowledge construction are almost exclusively limited to the domain of e-learning by students at educational institutions. By demonstrating the effects of knowledge construction in a non-educational domain, this study confirms the relevance of knowledge construction beyond the domain of e-learning at educational institutions.

1.3.2 Significance to Practice

Both active contributors to asynchronous online discussions and lurkers (legitimate peripheral participants—participants who read discussion posts but rarely contribute, see section 2.4 for a review of literature about lurkers) are exposed to transcripts of the discussions, and thus are consumers of the information generated in the discussions. Particularly for lurkers, who tend to be the majority, taking part in a discussion is essentially limited to being exposed to the transcript.

For health-related online discussions, actions of the consumers of such information based on the information can have consequences for their own as well as others' wellbeing, and, ultimately, for the good of the society at large. Therefore, it is of practical interest and importance to investigate if scrutiny and clarification of recommendations made over an online discussion (achieved via knowledge construction activities in the discussion) result in better quality information.

1.4 Definition of Terms

Online community. A group of people exchanging messages at an Internet-based forum on a sustained basis. Members of an online community are tied together because of the similarity of their objectives, interests, and needs, but, in most cases, not by geographical proximity. Active participation may result in emotional bonds between members. Online communities are self-managing; they develop their own policies and standards of behaviour (Preece & Maloney-Krichmar, 2003; Whittaker, Isaacs, & O'Day, 1997).

Online discussion. A discussion conducted over the Internet via discussion forum software to address a particular topic, question, or issue (Gunawardena, Lowe, & Anderson, 1997; Moore & Marra, 2005).

Discussion thread. A captured sequence of messages exchanged in an online discussion. Discussion forum software may group such messages hierarchically for easy reading.

Contributor or **poster**. An individual who posted a message to a discussion forum, either starting a new discussion or contributing to an ongoing discussion.

Lurker. An individual that reads discussion forum messages, but never posts any messages.

Abductive inference. According to Krippendorff (2004), in context of content analysis abductive inference refers to interpretation from the results of content analysis in relation to contextual phenomena (multiple interpretations may be possible, so that researchers have to rely on their judgment in making such an interpretation). In this research, interpretation from the results of content analysis of discussion thread messages in relation to the context in which the messages were created and in relation to the context in which they are consumed. This research did not study these contexts empirically, as it was limited to studying publicly available content.

More broadly, abductive inference refers to "inference to the best explanation" (Douven, 2011)—to inferring explanations from secondary data. This type of inference is frequently used both in everyday life and in scientific reasoning, and people often use

abductive reasoning without being aware of the term. I explicitly used the term "abductive inference" in the present study to keep the terminology consistent with Krippendoff (2004), which was the main source of the methodology used in the study.

1.5 Theoretical Framework

This study was based on the social constructivist theory of learning and on the connectionist view of KM.

Social constructivism suggests that learning happens via co-construction of meanings in social interactions (Palincsar, 1998; Vygotsky, 1978). Learners form tentative mental models, which they test by externalizing them to get feedback from others (other learners and experts). Comments result in revisions to the mental models, with the knowledge improved both at the individual level and at the group level. Knowledge construction in a group occurs via externalization, exchange, revision, and re-internalization of individual knowledge.

Connectionist view of KM suggests that knowledge resides both in individuals and in networks of individuals (such as the individuals in communities of practice) participating in knowledge sharing and construction (Dalkir, 2011). Knowledge is not only captured from the organization's environment, but is also developed internally. As individuals belonging to a network interact and share perspectives, new knowledge is created, which can partially be captured in documents. Knowledge in the heads of individuals belonging to the network is tacit knowledge. Knowledge captured in documents is explicit knowledge. The distinction between tacit and explicit knowledge was initially introduced by Nonaka and Takeuchi (1995).

In terms of tacit and explicit knowledge, I depict the dynamics of discussion forum discussions as follows. Documents both contain the captured knowledge (reflecting the tacit knowledge of the individuals) and reflect the processes that led to knowledge construction. In online discussions interactions between individuals result in updates to their tacit knowledge; at the same time, the interactions are captured as a discussion thread forming a transcript of the discussion (see Figure 1). In transcripts of online discussions the knowledge construction processes are captured in considerable detail.

Thus, by analysing discussion transcripts (explicit knowledge) one can make abductive inferences regarding the knowledge in the network and the processes of knowledge construction (for an explanation of the term "abductive inference", refer to section 1.4).

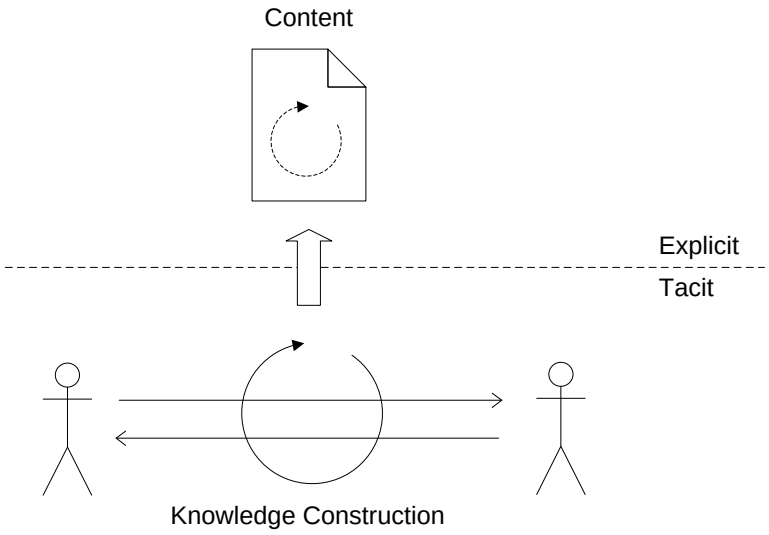

Figure 1. Knowledge construction in an online discussion.

Based on the social constructivist theory of learning and on the connectionist view of KM, this study takes a view that knowledge is constructed in health support group online discussions. The results of the analysis of discussion content allow one to make meaningful abductive inferences regarding the knowledge construction process and its outcomes with respect to the tacit knowledge of the discussion group.

1.6 High-Level Hypotheses

This study conceptualizes the knowledge construction process reflected in transcripts of online discussions as a single construct: level of knowledge construction (in this manuscript often abbreviated as just "knowledge construction"). Level of knowledge construction is the extent to which the interactions captured in the transcript can be seen as interactions leading to the construction of knowledge.

The outcomes of knowledge construction are considered from two perspectives. First, the construct of perceived information quality reflects the perspective of individuals drawing knowledge from the discussions (see section 2.8 for a review of literature on

perceived information quality). It is assumed that relevant, useful, and easy to understand transcripts correspond to discussions that lead to updates in the tacit knowledge.

Second, the construct of information integrity reflects the perspective of experts—whether the information is objectively valid from the point of view of the current state of scientific knowledge (see section 2.9 for a review of literature on information integrity). It is assumed that transcripts containing knowledge that is objectively valid correspond to discussions in which the updates to tacit knowledge are of high quality. (Even though tacit knowledge is contextualized and subjective knowledge, it is assumed that if it is based on objectively validated general knowledge, it is likely to lead to better results when applied.) In the context of health support group online discussions, information integrity is seen as the extent to which the suggestions for health management actions made in course of the discussions are consistent with evidence-based medicine.

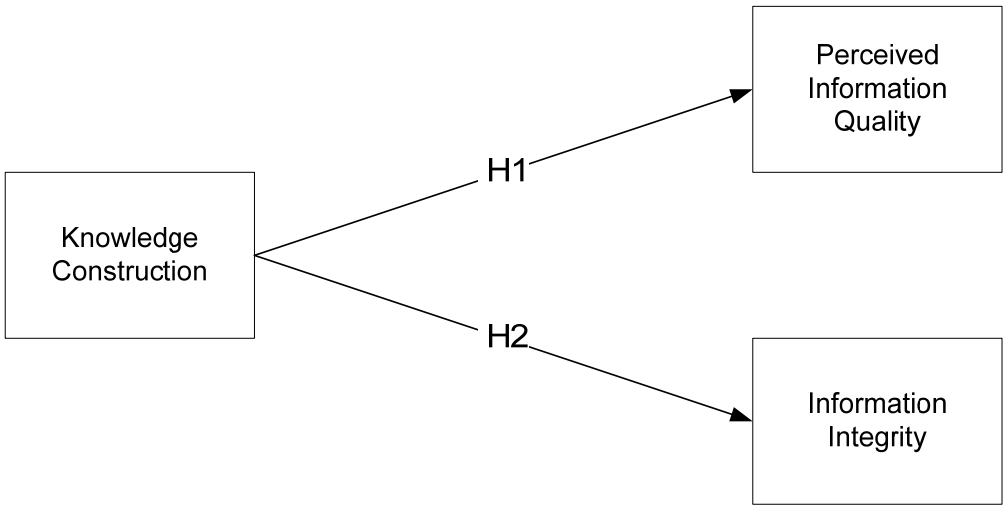

Figure 2. High-level research model.

I formulate the research questions as high-level hypotheses (see Figure 2). These hypotheses are restated at a lower level in section 3.3.

H1: Knowledge construction results in higher perceived information quality.

Discussion forum contributions are presented to the discussion participants as information accessed online. Therefore, it is meaningful to consider discussions from the perspective of the quality of such information. Moreover, an important category of discussion forum participants are lurkers, for whom the effect of the discussion is, arguably, limited to the effect of such information, which they consume similarly to any other online document. Higher quality discussions, involving knowledge construction activities, have been shown to result in better quality knowledge gained by the participants (Chiu, Hsu, & Wang, 2006). I take the view that some of this knowledge gain is due to higher quality information (from the perspective of the information consumers) generated in an insightful discussion involving knowledge construction. Thus, in my view, explicitation (exploring different perspectives and providing clarifications) and evaluation (re-examining the credibility) happening in a natural dialogue, when captured, result in information perceived by its consumers to be of higher quality compared to discussions that do not involve knowledge construction (and thus, are limited to stating facts without clarification and re-examination). For a detailed introduction of the concepts of explicitation and evaluation, refer to section 3.2.1.

H2: Knowledge construction activities in health support group online discussions result in greater prevalence of evidence-based knowledge.

Not all information shared in health support group online discussion is reliable from medical/clinical perspective. A prior study of cancer support group online discussion reported that some of the information shared was misleading, and that such misleading information was often identified and corrected in the course of the discussions (Esquivel, Meric-Bernstam, & Bernstam, 2006). Thus, one would expect that discussions involving knowledge construction (and thus, critical re-examination of contributions) would result in evidence-based content being identified and in anecdotal content being rejected, as the participants identify the sources of knowledge and evaluate their credibility. Thus, in my view, discussions involving knowledge construction should result in greater prevalence of evidence-based knowledge.

1.7 Summary of Methods

The hypotheses formulated in section 1.6 were tested by using content analysis of the publicly available content of health support group online discussions. The unit of analysis was a discussion thread. Thus, content analysis was used to measure latent variables of knowledge construction, perceived information quality, and information integrity for a sample of randomly selected threads. Then, statistical analysis was used to test the hypotheses, using a combination of PLS SEM and linear regression techniques.

Existing instruments were adapted for knowledge construction and perceived information quality. For information integrity, a simplified Delphi technique was used (facts were extracted from the content by coders and were assessed by domain experts).

When interpreting the outcomes of hypothesis testing, abductive inferences were made to the contexts of generation and use of the content.

1.8 Organization of the Thesis

The rest of this dissertation is organized as follows.

Chapter 2 reviews the literature relevant to the problem addressed by this study. First, the chapter introduces the approach used to conduct the literature review. Then, the chapter focuses on describing the practice of health support group online discussions in context of the shift to patient centred healthcare. Then, social constructivism as the theoretical foundation of this study is briefly introduced, followed by a detailed review and analysis of the existing research involving coding transcripts of online discussions to detect and to measure knowledge construction (overwhelmingly, these studies have been conducted in context of e-learning). The chapter continues by reviewing the studies of knowledge construction and online discussions conducted from a knowledge management perspective. The chapter concludes by discussing the existing conceptualizations and measures of perceived information quality and information integrity and by summarizing the knowledge gaps discovered in the literature.

Chapter 3 introduces in detail the conceptualizations of knowledge construction, perceived information quality, and information integrity. Then, the research model is introduced, along with detailed justifications of the research hypotheses.

Chapter 4 describes the research methods used in this study. First, the overall approach to research is stated as positivist content analysis. Then, the approach to content analysis employed in this research is characterized from the perspective of the criteria suggested by Krippendorff (2004). Then, the choice of the content source and this study's approaches to unitizing, sampling, screening, and coding are discussed. Then, the operationalisations of the constructs are introduced. The chapter continues by discussing the steps taken to assess and to ensure reliability and validity and concludes by introducing the statistical techniques used for hypotheses testing.

Chapter 5 presents the results of data analysis. Descriptive statistics and inter-rater reliability are discussed, followed by the results of a preliminary test of convergent and discriminant validity using exploratory factor analysis. Then, the results of measurement and structural model testing are introduced. The chapter continues by presenting a range of post-hoc analyses conducted to further explore the meaning of the results, involving the analysis of variations of the research model.

Chapter 6 summarizes the results and discusses their implications. A summary of the findings is followed by a discussion of the findings from the perspectives of theory and practice. The chapter ends by discussing the limitations of the research, by making suggestions for future research, and by briefly restating the conclusions for the whole thesis.

Chapter 2. Literature Review

2.1 Introduction

This chapter reviews the literature and states the knowledge gap addressed by this study. First, the chapter describes the method used to conduct the literature review. Then, the practice of health support group online discussions is discussed in the context of the ongoing shift to patient centred healthcare. Moreover, lurkers (legitimate peripheral participants) are introduced as important stakeholders in online discussions.

Then, the chapter introduces social constructivism as the theoretical foundation of this study and reviews the existing studies involving coding transcripts of online discussions to detect and to measure knowledge construction, with the focus on the coding schemes employed (overwhelmingly, these studies have been conducted in context of e-learning). The operationalisations of knowledge construction and the approaches to content analysis used in these studies are systematically compared.

The chapter continues by reviewing the studies of knowledge construction and online discussions conducted from a knowledge management perspective. A descriptive empirical study of online discussions at a community of practice is discussed in detail.

Then, the chapter introduces perceived information quality and information integrity as the main dimensions of information quality identified in the literature.

The chapter concludes by summarizing the knowledge gaps discovered. In addition, to cover background material highly relevant to understanding the scope and the implications of the outcomes of the present study (covered in Chapter 6), a section reviewing the literature on health literacy is included.

2.2 Method Used to Conduct the Literature Review

To conduct the literature review, I followed the approach suggested by Webster and Watson (2002). First, I searched online databases for combinations of keywords suggested by the purpose (see section 1.2) and by the high-level hypotheses (see section

1.6) of the present study. Then, I considered studies cited in the studies found to be relevant and articles citing the studies found to be relevant. I repeated these steps iteratively, until no new relevant articles could be found.

The databases were Google Scholar (http://scholar.google.com) and Thomson Reuters Web of Science (http://isiknowledge.com). I used keywords such as "knowledge construction", "online discussion", and "information quality". The purpose and the research questions of the present study did not imply the necessity of conducting a systematic review that could be repeated by other researchers. Therefore, I did not document all variations of keywords that I have used, and I did not use an independent coder to assess the reliability of judgments regarding the relevance of particular articles to the present study.

As suggested by Webster and Watson (2002), I used a concept matrix to organize the literature relating to the most important and the richest theme discovered in the literature review—the theme of conceptualization and measurement of knowledge construction in online discussions (see Table 2-18 in section 2.6.6.1).

Some approaches to reviewing literature alternative to the approach by Webster and Watson (2002) are the ones suggested by Khan, Kunz, Kleijnen, and Antes (2003) and Greenhalgh et al. (2005). The approach by Khan et al. (2003) involves five steps: (1) framing questions for a review, (2) identifying relevant work, (3) assessing the quality of the studies, (4) summarizing the evidence, and (5) interpreting the findings. The approach that was used in the present study can be described in terms of the approach by Khan et al. As to stage (1) in the approach by Khan et al., questions for the review were implied by the purpose and the high-level hypotheses of the study. As to stage (2), the search was not limited to searching by keywords and to scanning tables of contents, but I also followed the citation tree; however, I did not set rigid formal criteria for inclusion and relied on my own judgment (because conducting a formal systematic review was not suggested by the purpose of the present study). As to stage (3), I did assess the quality of the studies (and commented on quality aspects such as inter-rater reliability in the review). For stage (4), I summarized the evidence focusing in particular on conceptualization and operationalization of knowledge construction (see

section 2.6.6). Finally, for stage (5), I interpreted the findings by stating the knowledge gap to be addressed in the present study (see section 2.10).

The approach to conducting a literature review by Greenhalgh et al. (2005) suggests the use of a multidisciplinary research team and relies on interaction with the client for which the review is conducted. Clearly, the resource constraints of the present study did not allow to involve a multidisciplinary research team. Moreover, there was no client (the literature review in the present study was not conducted for a specific entity). Yet, it is of interest to compare the approach by Greenhalgh et al. to the approach used in the present study, to highlight the limitations of the approach used in the present study and to suggest how a broader and more formal and systematic review could be conducted. The approach by Greenhalgh et al. is intended to cover a broad research topic rather than to support the research focused on addressing a specific research problem and hypotheses. In particular, the approach by Greenhalgh et al. suggests emphasizing the history of how different research paradigms are applied in addressing the topic and emphasizing the comparison of how different research paradigms are applied to address the important dimensions of the focal topic or problem. Clearly, in future research it would be of much value to review the broader topic of knowledge construction by using the approach by Greenhalgh et al. Such research would require high levels of funding to form a multidisciplinary team that represents different research paradigms, with sufficient redundancy to allow formal inter-rater reliability assessment.

To assess the actual use of the three approaches to conducting a literature review introduced in this section, I compared the number of citations of the respective articles based on Google Scholar searches. The results are summarized in Table 2-1.

As seen in Table 2-1, the approach by Webster and Watson (2002) has been cited more frequently and is the only of the three approaches considered that has been cited in the two flagship MIS journals: MIS Quarterly and Information System Research. Therefore, in relying on the approach by Webster and Watson I followed the established common practice.

Table 2-1 Numbers of Citations of Articles Introducing Approaches to Conducting a Literature Review

Paper introducing an approach to conducting a literature review	Numbers of citations since 2008		
	All	In MIS Quarterly	In Information Systems Research
Webster and Watson (2002)	1024	4	7
Khan et al. (2003)	43	0	0
Greenhalgh et al. (2005)	175	0	0

Note. The numbers are based on Google Scholar searches conducted on 25 March, 2013.

2.3 Patient-Centred Care and Online Health Support Groups

Patient-centred care is a new paradigm in healthcare that promotes the involvement of patients themselves and their families in dealing with the patients' health issues (Robb & Seddon, 2006). Patient-centred care has been defined in a number of ways, reflecting slightly different perspectives (Mead & Bower, 2000). Some examples of such definitions are given in the following paragraphs.

According to Balint (1969), a patient-centred approach is an approach that views each patient as a unique human being. A patient should not be viewed simply as a case of the patient's specific illness, but rather seen as a whole human being.

According to Byrne (1976), patient-centred care is an approach to consultation taking into account the patient's knowledge and experience with respect to his or her own condition. Thus, each patient is an expert on his or her own condition.

A number of authors defined patient-centred care as an approach in which the physician views the illness through the patient's eyes and involves the patient in decision making regarding his or her illness (Grol, De Maeseneer, Whitfield, & Mokkink, 1990; Lipkin, Quill, & Napodano, 1984; McWhinney, 1985; Winefield, Murrell, Clifford, & Farmer, 1996).

According to Laine and Davidoff (1996), patient-centred care suggests care closely aligned with and responsive to the patients' particular needs and preferences.

These definitions suggest that the patient's knowledge and experiences play a particularly important role in patient centred care.

Patient-centred care views both the patient and the doctor as persons. The doctor treats the patient in a respectful manner and shares information with the patient. In term of expertise, the patient and the doctor are not equal. Nonetheless, the patient is seen as having knowledge that is relevant to decision making regarding his or her condition. Sharing of information, communication, and decision making in patient-centred care are based on person-to-person relationships between patients and their doctors (Duggan, Geller, Cooper, & Beach, 2006).

In patient-centred care, the patient becomes increasingly knowledgeable about his or her condition (Lipkin et al., 1984). The patient gains knowledge from consultations with the doctor. Sharing complete and unbiased information with patients and their families is a part of knowledge enabling in the patient-centred care paradigm. The patient and their family receive timely, complete, and accurate information in order to effectively participate in care decision-making. Moreover, in patient-centred care doctors encourage patients to gain knowledge not only from direct consultations, but also from other sources, such as from their peers, from books, from periodicals, or online.

In line with the patient-centred care paradigm, people are becoming more active at finding information related to their health (Suggs, 2006). They use Internet to access health databases and join online health support groups to discuss specific issues with their peers. People are becoming more knowledgeable about healthcare and medicine because of the wide access to health-related information via the Internet.

In an opinion article, Wright and Bell (2003) summarized the benefits of peer-to-peer Internet-based health support groups claimed in the literature. The main benefits were the ability to gain access to diverse information and diverse points of view, the reduced stigma of illness because of moral support from people with similar problems, and the ability to disclose personal information safely because of anonymity. The negative aspects were also highlighted, including the presence of hostile messages and the difficulty in forming long-term relationships in the electronic environment.

Eysenbach, Powell, Englesakis, Rizo, and Stern (2004) reviewed the empirical studies of the effects of peer-to-peer Internet-based support groups as medical interventions. Based on their review of 76 studies using a variety of experimental designs, they concluded that there was no robust evidence of the effects of peer-to-peer Internet-based support groups because in most cases the differences in desired outcomes between the experimental group and the control group were not statistically significant. Yet, among the studies that did find a statistically significant difference, the difference was in almost all cases in favour of the experimental group (the group that was involved in online discussions); and in some of the studies a statistically significant correlation was found between the extent of involvement in peer-to-peer support groups and the desired health outcomes (although the direction of the cause-effect relationship was not clear).

Eysenbach et al. (2004) indicated that the lack of robust evidence was, most likely, because of the low power of the studies. Too few participants were involved, and the participants who had access to peer-to-peer discussions often did not use them. In my view, this underlines the difficulty of using experiments to study online health support groups. An effective online support group emerges spontaneously and cannot be administered in a controlled manner, such as can be a medical drug.

Lasker, Sogolow, and Sharim (2005) investigated the motivation of people with rare diseases to join health support group online discussions and the types of contributions by people at different stages of their illness. Lasker et al. used content analysis with a sentence as a unit of analysis. The coding scheme included three dimensions (all measured on categorical scales): biomedical—health condition discussed in the content, socioemotional—emotional state expressed in the content, and organizational—organizational and financial issues addressed in the content. Two coders analysed 710 messages (the number of sentences in these messages was not reported) generated by 275 contributors during two months of discussions at a health support group devoted to primary biliary cirrhosis. Coder background and coder training were not discussed. Inter-rater reliability was reported only for codes after negotiation between coders; the details of the coder negation protocol were not reported.

Based on the analysis of the prevalence of different types of topics, Lasker, Sogolow, and Sharim (2005) concluded that the health support group had a biomedical rather than a socioemotional or organizational emphasis. In other words, the members of the community took part to obtain health-related information, rather than to seek moral support or to resolve organizational or financial issues.

2.4 Lurkers

Katz (1998) quoted a survey by a consulting firm in Chicago that found that 98% of discussion group members at AOL (http://www.aol.com) and MSN (http://www.msn.com) were lurkers. Even though early studies viewed lurkers as free riders, later studies emphasized positive roles of lurkers in online communities (Nonnecke & Preece, 2000).

Takahashi, Fujimoto, and Yamasaki (2003) identified in their research the so called "active lurkers"—lurkers who ultimately contribute to the effectiveness of online communities in indirect ways. The study categorized active lurkers as propagators who share the knowledge that they gain from the online communities' discussion groups with a broader community and as practitioners who apply this knowledge in their family and work contexts. Furthermore, the study emphasized the importance of active lurkers in promoting their community to the outside.

Because the term lurker can be perceived as having negative connotations, Nonnecke and Preece (2000) argued that the term does not reflect the positive outcomes of the behaviour it denotes and proposed to use instead the term "legitimate peripheral participants". Legitimate peripheral participants enhance the learning resulting from the online discussion—the legitimate peripheral participants themselves learn from the discussion, and the legitimate peripheral participants (as active lurkers) spread the knowledge outside the immediate online community. The form of participation in online communities exercised by legitimate peripheral participants should be seen as normal and constructive.

2.5 Constructivism and Knowledge Construction

Social constructivism is the main theoretical foundation of this study. This section introduces social constructivism. Individual constructivism is introduced first; then, social constructivism is introduced and contrasted with individual constructivism.

Here, the term *individual constructivism* is used to refer to constructivism as it was introduced by Piaget (Lefrançois, 1995), to distinguish it from *social constructivism* introduced by Vygotsky (Palincsar, 1998; Vygotsky, 1978).

2.5.1 Individual Constructivism

According to individual constructivism, individuals obtain new knowledge by constructing it based on their existing knowledge and on experiences gained as they interact with their environment (Lefrançois, 1995). As an individual responds to situations and experiences the outcomes, his or her ideas about the world change and develop, and this is seen as knowledge construction. Thus, according to individual constructivism, different individuals can have different understanding of reality; differences in constructions between individuals are legitimate and expected, because individuals construct knowledge, rather than discover objective knowledge.

2.5.2 Social Constructivism

Comparing to individual constructivism, social constructivism emphasizes the special role of social interactions and shifts the focus from knowledge construction by individuals to knowledge construction by groups of interacting individuals.

According to social constructivism (Palincsar, 1998; Vygotsky, 1978), knowledge is constructed in social interactions, as individuals negotiate meanings. Language plays a central role in learning because it mediates external experiences into shareable and negotiable understanding.

Social constructivism has been highly influential in education—it has changed the way educational institutions view the learning process (Palincsar, 1998). Modern educational institutions encourage students to become active learners by

communicating their ideas or problems to others and by using discussions to solve problems via gaining multiple perspectives. Discussions are viewed as promoting higher levels of understanding because of the need to state, to restate, and to develop propositions to clarify, to criticize, and to make arguments.

2.6 Conceptualizations and Measures of Knowledge Construction in Online Discussions

According to social constructivism, learning is facilitated when individuals interact. Such interactions may be enabled by technology, such as computer supported collaboration learning (CSCL) environments (Alavi, 1994) or other environments enabling online discussions. Although numerous studies have considered knowledge construction in discussions in e-learning (De Wever, Schellens, Valcke, & Van Keer, 2006), there are few empirical studies of knowledge construction in computer mediated discussions in non-educational contexts. The focus of the present study is on health support group online discussions—a context clearly different from e-learning. Even though pedagogical strategies used in e-learning to promote knowledge construction are not relevant to this study, measures of knowledge construction in online discussions developed in the context of e-learning may be relevant.

The existing measures of knowledge construction that can be used in content analysis are discussed in this section. All of these studies have relied on social constructivism (introduced in section 2.5.2) as the primary theoretical foundation. The section starts by focusing on the measures that have been used in multiple studies and thus are better validated. Then, the rest of the measures are discussed. The section concludes by comparing the existing measures of knowledge construction from measure content and measure validity perspectives.

The following subsections (up to and including the subsections of section 2.6.5) are named based on the declared aims of the studies covered.

2.6.1 Henri (1992)—Learning Process Framework

Henri (1992) developed five dimensions to describe the process of learning in CSCL environments, which were Participation, Social, Interactive, Cognitive, and Metacognitive. The Henri's scales for the Participation and Interactive dimensions are not discussed here because they are not relevant to knowledge construction. The details of the Social, Cognitive, and Metacognitive dimensions are given in Table 2-2 (The Social dimension did not describe knowledge construction, but played an important supplementary role in the analysis of knowledge construction). Henri suggested using a sentence as a unit of analysis, although his scale is applicable at other levels, such using a message as a unit of analysis.

Table 2-2 *Henri's (1992) Dimensions of Learning*

Dimension	Subdimension	Scale	Levels	Categories
Social		Nominal		*related to the subject, unrelated to the subject*
Cognitive[a]		Ordinal	*elementary clarification, in-depth clarification, inference, judgment, strategy*	
Metacognitive[a]	Knowledge	Categorical		*assessment of the task, learner self-perception, means to succeed*
Metacognitive[a]	Skill	Categorical		*evaluation, planning, regulation, self-awareness*

[a]Assessed only for units related to the subject of discussion.

The Social dimension reflected if the content of the unit of analysis was related to the intended subject of the discussion. The Cognitive and Metacognitive dimensions were assessed only for units related to the subject of the discussion (and did not make sense for units that were purely social in nature).

The Cognitive dimension reflected the level of critical thinking exhibited in a sentence by its contributor and was measured on an ordinal scale, thus allowing comparison of the levels of critical thinking between units.

The Metacognitive dimension reflected awareness, self-control, and self-regulation of learning. The subdimensions of Metacognitive, Knowledge and Skill, were measured on categorical scales.

The known uses of the measure developed by Henri (1992) are discussed in the following subsections.

2.6.1.1 Bullen (1998)—Participation and Critical Thinking

Bullen (1998) adapted the Henri's (1992) model to investigate the use of computer-mediated communication (CMC) to support the process of learning. Bullen aimed to determine whether the students were actively building on each other's contributions and thinking critically about the discussion topics.

Bullen (1998) studied synchronous online discussions in a computer information systems class.

Table 2-3 *Bullen's (1998) Scale for Critical Thinking*

Dimension	Scale	Levels
Critical Thinking[a]	Ordinal	*low, moderate, high*

[a]Measured using an adaptation of the Henri's (1992) scale for cognitive.

Bullen (1998) adapted a subset of the Henri's (1992) scale (the Cognitive dimension only) to measure the level of critical thinking (see Table 2-3). The data were analysed as qualitative data, with the level of the Critical Thinking dimension used as a descriptive measure informing the qualitative analysis. Inter-rater reliability was not assessed.

Bullen (1998) found evidence of critical thinking.

2.6.1.2 Hara, Bonk, and Angeli (2000)—Knowledge Construction and Interaction

Hara, Bonk, and Angeli (2000) used the Henri's model to analyse interactions between students and teachers in online discussions. Hara et al. focused on the following aspects of the interactions: (a) student participation rates, (b) electronic interaction rates, (c)

social cues within student messages, (d) evidence of cognitive and metacognitive processing in student messages, and (e) the depth of processing within student messages.

Content analysis was used for (c) and (d): Hara et al. (2000) used the Henri's (1992) scales for Social, Cognitive, and Metacognitive. Hara et al. found it difficult to use the Knowledge sub dimension of Metacognitive together with Cognitive because they found them to be similar. Therefore, they used only the Skill sub dimension of Metacognitive in their final analysis.

Hara et al. (2000) analysed 271 messages collected over 15 weeks of discussion by 20 students enrolled in a graduate Psychology course at a university in the USA. A paragraph was used as a unit of analysis (the number of paragraphs in the messages analysed was not reported).

Independent researchers were used as coders. It was not reported if the coders were trained. The coders negotiated the final score, but the details of the negotiation protocol were not reported. Inter-rater reliability was assessed for the initial scores (the scores before negotiation) using percent agreement; the details of inter-rater reliability results are summarized in Table 2-4.

Table 2-4 Inter-Rater Reliability in Hara et al. (2000)

Measured by	Social	Cognitive	Metacognitive / Skill
Percent agreement	0.78	0.75	0.71

Percent agreement has been criticized as a measure that overestimates inter-rater agreement (Riffe, Lacy, Fico, & Fico, 2005), which is particularly true for nominal scales. Hara et al. (2000) used a nominal scale for Social, and the scales for Cognitive and Skill subdimensions of Metacognitive involved, respectively, five levels and three categories. For a scale involving three choices, chance agreements are less likely than for a nominal scale (so that random sequences result in 33.33% agreement compared to 50% for a nominal scale). For a scale involving five choices, chance agreements are even less likely. Thus, one can judge the inter-rater reliability achieved by Hara et al.

for Cognitive and Metacognitive as relatively good, even though it was below the 0.8 threshold suggested by Riffe et al. (2005).

Based on a qualitative analysis of messages (informed by the outcomes of the content analysis), Hara et al. (2000) concluded that online discussions promote cognitive processes.

2.6.1.3 Pena-Shaff & Nichols (2004)—Knowledge Construction and Participation

Table 2-5 Categories Retained by Pena-Shaff and Nicholls'(2004) That Matched Levels and Categories in the Henri's (1992) Model

Pena-Shaff & Nicholl (2004)	Henri (1992)	
Category	Category or level	Dimension / subdimension
clarification	*elementary clarification and in depth clarification*	Cognitive
interpretation	*inference*	Cognitive
judgment	*judgment*	Cognitive
reflection	*self-awareness*	Metacognitive / Skill
reply	*direct responses and direct commentaries*	Interactive
other	*unrelated with the subject*	Social

Pena-Shaff and Nicholls (2004) studied student interactions and meaning construction in online bulletin board discussions. Pena-Shaff and Nicholls developed 11 knowledge construction categories based on qualitative analysis of the messages. Qualitative analysis started by using a coding scheme that included categories and levels from the Henri's (1992) model (along with categories from other related work), which were simplified for the purposes of the qualitative analysis. The coding scheme was allowed

to evolve inductively: New categories were allowed to emerge in the course of the analysis, and not all of the initial categories were retained.

Pena-Shaff and Nicholls (2004) analysed 152 messages in discussion threads randomly sampled from discussions by 35 students enrolled in an advanced communication course at a university in the USA. A sentence was used as a unit of analysis.

Many of the categories retained after the analysis was completed matched the categories and levels of the dimensions in the Henri's (1992) model. The categories in the Pena-Shaff and Nicholls' (2004) analysis that matched the categories or levels of the dimensions in the Henri's model relevant to knowledge construction are given in Table 2-5.

Pena-Shaff and Nicholls (2004) found the highest proportions of messages in the *clarification* (44%) and *interpretation* categories (15%), which matched low and middle levels of the Cognitive dimension in the Henri's (1992) model.

Because Pena-Shaff and Nicholls (2004) used qualitative analysis, inter-rater reliabilities did not apply (and were not reported).

2.6.2 Gunawardena, Lowe, and Anderson (1997)—Knowledge Construction in CMC Environments

Gunawardena, Lowe, and Anderson (1997) developed a coding scheme for the analysis of the process of knowledge construction in CMC environments. Gunawardena et al. used the grounded theory to develop their coding scheme; thus, the coding scheme was developed inductively, based on a qualitative analysis of captured discussions rather than on prior theory.

Gunawardena, Lowe, and Anderson (1997) analysed 206 messages generated in one week of online debates by graduate students (at a university in the USA) who participated in an online conference devoted to CMC issues in distance education. A message was used as a unit of analysis.

The codes that resulted from inductive qualitative analysis of the messages suggested five phases of knowledge construction, including: (1) sharing/comparing information,

(2) exploration of dissonance among ideas, (3) negotiation of meaning, (4) testing and modifying proposed synthesis, and (5) achieving agreement (See Table 2-6). Thus, implicitly, Gunawardena et al. suggested a single dimension to measure the process of learning, with five levels corresponding to the phases. These phases are congruent with the levels of learning in the Henri's (1992) scale for Cognitive. Thus, higher phases in the Gunawardena et al. scale can be seen as corresponding to higher levels of knowledge construction.

Table 2-6 Gunawardena et al.'s Phases of Knowledge Construction

Dimension	Scale	Levels
Phase of Knowledge Construction	Ordinal	*1-sharing/comparing information, 2-exploration of dissonance among ideas, 3-negotiation of meaning, 4-testing and modifying proposed synthesis, 5-achieving agreement*

Sharing/comparing information (Phase 1) referred to new information or new discussion topics appearing in the discussion. *Exploration of dissonance among ideas* (Phase 2) referred to the recognition of inconsistencies between statements made in Phase 1 and the contributors' existing knowledge. *Negotiation of meaning* (Phase 3) referred to the resolution of dissonances discovered in Phase 2 via redefinition of terms or via adjustments of the topic. *Testing and modifying proposed synthesis* (Phase 4) referred to the re-examination of meanings tentatively established in Phase 3. *Statement of agreement* (Phase 5) referred to the establishment of contributors' agreement regarding the synthesis.

The distribution of messages to phases was as follows: sharing/comparing information (Phase 1)—191 postings, exploration of dissonance among ideas (Phase 2)—5 postings, negotiation of meaning (Phase 3)—4 postings, testing and modification proposed synthesis (Phase 4)—2 postings and statement of agreement (Phase 5)—4 postings. Thus, the overwhelming majority of the messages were in Phase 1.

Gunawardena et al. (1997) concluded that informal interaction is not congruent with the process of knowledge construction (does not promote and possibly prevents knowledge construction).

Because Gunawardena et al. (1997) used qualitative analysis, inter-rater reliability did not apply (and was not reported).

The only known use of the measure developed by Gunawardena et al. (1997) in subsequent research is discussed in the following subsection.

2.6.2.1 Kanuka and Anderson (2007)—Replication of the Study by Gunawardena et al.

Kanuka and Anderson (2007) investigated the process of knowledge construction in an online discussion by professionals. A survey questionnaire, telephone interviews, grounded theory, and content analysis were used in the Kanuka and Anderson study. Content analysis was conducted by using the coding scheme from Gunawardena et al. (1997).

Kanuka and Anderson (2007) analysed 252 messages collected over three weeks of discussions at a discussion forum devoted to the use of learning technologies at workplace learning centres. The messages had been posted by 25 managers located across Canada. A message was used as a unit of analysis.

Kanuka and Anderson (2007) assumed that the scheme by Gunawardenaet al. (1997) is useful in analysing the process of knowledge construction. However, they found that higher levels of learning did not occur in the discussion. In my view, this was for the following reasons.

First, discussions of professionals differ from discussions of learners in education. Professionals are likely to be satisfied by obtaining the information they seek and clarifying it, and therefore are less likely than students to engage in negotiation of meaning or consensus building. Moreover, unlike in educational discussions, in discussions among professionals there is no encouragement to engage in higher levels of knowledge construction; indeed, professionals might view such discussions as not the best use of their time (and thus, there might be peer pressure or pressure from the managers to discourage discussions going beyond addressing specific information needs at hand). In addition, the style of writing required to reach higher phases of knowledge construction (as defined by Gunawardena et al., 1997) may be seen as not

appropriate for workplace communication, and the contributors might not be skilful enough in using such style of writing even if they do not feel that it is discouraged.

Second, the coding scheme of Gunawardena et al. (1977) did not define the higher phases of knowledge construction clearly enough, resulting in an analysis biased in favour of lower phases of knowledge construction. It may be easier for a coder to identify new information than to identify the higher phases of knowledge construction.

Kanuka and Anderson (2007) did not report inter-rater reliability.

2.6.3 Veerman, Andriessen, and Kanselaar (1999)—Collaborative Learning and Argumentation

Veerman, Andriessen, and Kanselaar (1999) investigated the process of learning in synchronous and asynchronous CMC. The study involved three systems. Two of the systems enabled synchronous communication; one of them included a visual mapping tool (SynVM), and another did not have a visual mapping tool (Syn). The remaining system enabled asynchronous communication and did not include a visual mapping tool (Asyn).

Veerman et al. (1999) developed three dimensions to analyse the process of learning in CMC: Focus, Argumentation, and Constructive Activities. The Veerman et al.'s scale for Argumentation is not discussed here because it is not relevant to knowledge construction. The details of Focus and Constructive Activities dimensions are given in Table 2-7 (The Focus dimension did not describe knowledge construction, but played an important supplementary role in the analysis).

Table 2-7 *Veerman et al.'s (1999) Dimensions*

Dimension	Scale	Categories
Focus	Nominal	*task related, non-task related*
Constructive Activities	Categorical	*addition, evaluation, explanation, summaries, transformation*

The Focus dimension referred to whether the message content addressed the declared subject of the discussion. The Argumentation and Constructive Activities dimensions

were considered only for units of analysis rated as *collaborative learning*. The Constructive Activities dimension referred to activities that resulted in knowledge construction in the discussion.

Veerman et al. (1999) analysed 61 discussion threads generated over three weeks of discussions at an educational technology course in a university in the Netherlands (20 in Syn, 13 in SynVM, and 28 in Asyn). A message was used as a unit of analysis. The numbers of messages in the discussions were not reported.

The same set of messages was used to tune the coding scheme and to derive conclusions. Veerman et al. (1999) found higher proportion of task related messages in Asyn (88%), compared to 64% in Syn and 42% in SynVM.

Constructive activities were common in Asyn (72%), compared to 48% in SynVM and 35% in Syn. Constructive activities represented knowledge construction. This result suggests that asynchronous communication promotes knowledge construction better than synchronous communication.

Statistical tests for the significance of these differences were not conducted. Moreover, different systems were used by different groups of students, so that it is possible that some the differences were because of differences between groups rather than between systems.

The process used to recruit the coders was not reported. Presumably, the coders were the authors themselves. Coder training was not discussed. The inter-rater reliability is summarized in Table 2-8. High levels of inter-rater reliability are quite likely due to common understanding achieved between the authors as they worked together on the project.

Table 2-8 *Veerman et al.'s (1999) Inter-Rater Reliability*

Measured by	Focus	Argumentation	Constructive Activities
Cohen kappa (k)	0.91	0.89	0.74

The known uses of the measure by Veerman et al. (1999) are discussed in the following subsections.

2.6.3.1 Veerman and Veldhuis-Diermanse (2001)—Collaborative Learning in CMC

Veerman and Veldhuis-Diermanse (2001) investigated collaborative learning in CMC. Veerman and Veldhuis-Diermanse used four types of CMC. Two of the systems enabled synchronous communication: one included a visual mapping tool (SynVM) and another did not have a visual mapping tool (Syn). The remaining two systems enabled asynchronous communication and did not include a visual mapping tool; one of them allowed the users to explicitly label their messages as questions, answers and so on (AsynL), and another did not (Asyn).

Veerman and Veldhuis-Diermanse (2001) studied the process of knowledge construction in academic discussions. The coding scheme involved three dimensions, Collaborative Learning, Social, and a dimension categorizing messages not related to the topic of the discussion (not discussed further as not related to knowledge construction). The details of the Collaborative Learning and Social dimensions are given in Table 2-9 (The Social dimension did not describe knowledge construction, but played an important supplementary role in the analysis).

Table 2-9 *Veerman and Veldhuis-Diermanse's (2001) Dimensions*

Dimension	Scale	Levels	Categories
Collaborative Learning[a]	Ordinal	*new information, explanation, evaluation*	
Social	Nominal		*related to the subject, unrelated to the subject*

Note. Veerman and Veldhuis-Diermanse (2001) rather nonintuitively used the term Task-Related for the Collaborative Learning dimension. I renamed the dimension using the words Veerman and Veldhuis-Diermanse used to describe it. Veerman and Veldhuis-Diermanse did not include the Social dimension explicitly, but it was implied from their analysis. I included the dimension using the name for a similar dimension in the coding scheme by Henri (1992).

[a]Veerman and Veldhuis-Diermanse (2001) treated this scale as categorical. However, based on Schellens and Valcke (2005), I believe that this scale is ordinal, with *evaluation* corresponding to the highest level of knowledge construction and *new information*—to the lowest.

The Social dimension distinguished messages related to the intended topic of the discussion from the rest. The Collaborative Learning dimension categorized messages according to knowledge construction activities.

Veerman and Veldhuis-Diermanse (2001) analysed discussions by four different groups of undergraduate students enrolled in an educational technology course at a university in the Netherlands (different systems, Syn, SynVM, Asyn, and AsynL, were used by different groups). A message was used as a unit of analysis. The numbers of the discussions and messages analysed are listed in Table 2-10.

Table 2-10 The Numbers of Discussions and Messages Used for Analysis in Each Study

	Syn	SynVM	Asyn	AsynL
Number of analysed discussions	20	30	28	4
Number of analysed messages	2040	1287	952	1088

Veerman and Veldhuis-Diermanse (2001) found that there was a higher proportion of Collaborative Learning messages in the asynchronous (Asyn–88% and Syn–85%) compared to the synchronous (Syn–61% and SynVM–42%) discussions.

Inter-rater reliability was not reported. Because inter-rater reliability was assessed in the earlier related study by Veerman et al. (1999) involving some of the same authors, it is difficult to see why inter-rater reliability was not reported.

The process used to recruit the coders was not reported. Presumably, the coders were the authors themselves. Coder training was not discussed.

2.6.3.2 Schellens and Valcke (2005)—Effect of Cognitive Processing

Schellens and Valcke (2005) investigated the effects of asynchronous electronic discussions on cognitive processing. Schellens and Valcke used two coding schemes with the same data: the Veerman and Veldhuis-Diermanse (2001) scheme and the Gunawardena et al. (1997) scheme (see sections 2.6.3.1 and 2.6.2). They argued that the two schemes are related, with *new information*, *explicitation*, and *evaluation* in the Veerman and Veldhuis-Diermanse (2001) scheme corresponding, respectively, to *sharing*, *exploration*, and *negotiation of meaning* in the Gunawardena et al. (1997) scheme (see Table 2-6 in section 2.6.2).

Schellens and Valcke (2005) analysed 1428 messages posted over 4 months of discussions by 230 students enrolled in an undergraduate educational sciences course at a university in Belgium. The discussions were organized in four major topics, and the analyses were conducted separately for each topic. A message was used as a unit of analysis. The Veerman and Veldhuis-Diermanse (2001) scheme was used to distinguish task related messages, and the scheme of Gunawardena et al. (1997) was applied only to the task related messages.

Three independent researchers were used as coders. The coders were trained, but the details of the training protocol were not reported. The coders negotiated the final score, but the details of the negotiation protocol were not reported. Inter-rater reliability was assessed for the initial scores (before negotiation) using Cronbach's alpha; however, the details of inter-rater reliability were not reported for separate topics or by coding scheme. Rather, the article just mentions that inter-rater reliability ranged from 0.45 to 0.93 (presumably, across different topics and different coding schemes).

Collaborative learning messages dominated the discussion. In terms of the Veerman and Veldhuis-Diermanse (2001) scheme, the collaborative learning content was dominated by *new information* and *evaluation*—thus; both low and high levels of knowledge construction were well represented. In terms of the scheme by Gunawardena et al. (1997), the collaborative learning content was dominated by *sharing or comparing information* and by *negotiation of meaning* (as one would expect from the correspondence indicated in Figure 3); messages coded as *achieving agreement* or as *testing the proposed synthesis* were rare. However, values for correlations between the results for the two schemes were not reported.

A number of hypotheses relating to patterns of knowledge construction activities (such as a hypothesis that the level of knowledge construction grows over the course of a discussion) were tested in the study. However, the hypotheses were not based on a strong theoretical foundation, and, in my opinion, were primarily a form of presenting descriptive results.

In terms of the results of hypothesis testing, the results obtained by the two different coding schemes were consistent.

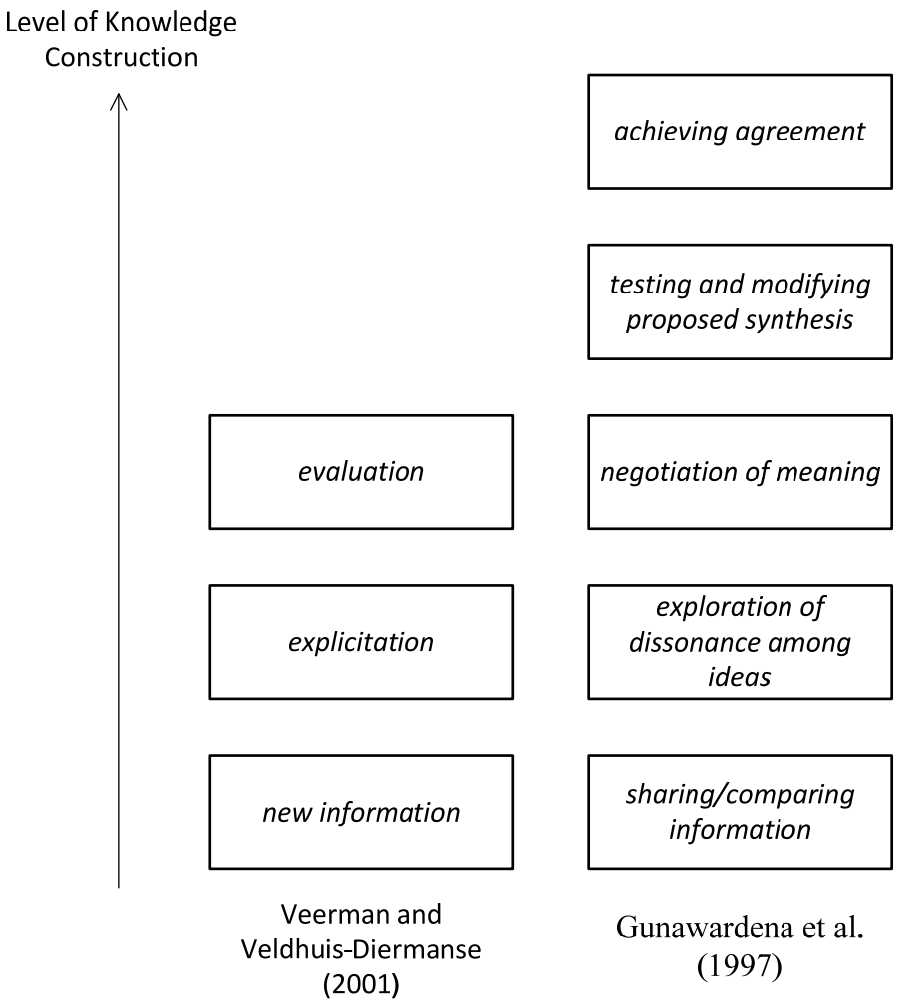

Figure 3. Mapping between Veerman and Veldhuis-Diermanse (2001) and Gunawardena et al. (1997) schemes.

To an extent, the consistency found between the two schemes may have been because the same coders rated all messages; the result may have reflected the coders' opinion about the two schemes. As researchers (even though independent), the coders were highly likely to have consciously or unconsciously formed an opinion, which may have resulted in bias. It is not clear from the article if the coders were able to communicate even before the formal negotiation; such communication would increase the likelihood of bias in the results.

2.6.3.3 Schellens and Valcke(2006)—Fostering Knowledge Construction

Schellens and Valcke (2006) investigated the effects of asynchronous electronic discussions on cognitive processing. The approach was similar to that of Schellens and

Valcke (2005). Two coding schemes were used with the same data: the Veerman and Veldhuis-Diermanse (2001) scheme and the Fahy et al. (2000) scheme. They argued that the two schemes are related, with *vertical questioning*, *horizontal questioning*, *statements*, and *reflection* in the Fahy et al. (2000) scheme corresponding to *new information* (*new facts*, *new experiences*, and *new theory*), *explicitation*, and *evaluation* in the Veerman and Veldhuis-Diermanse (2001) scheme (see Figure 4).

Schellens and Valcke (2006) analysed 1752 messages posted over 6 months of discussions by 300 students enrolled in an undergraduate instructional sciences course at a university in Belgium. The discussions were organized in six major topics, and the analyses were conducted separately for each topic. A message was used as a unit of analysis. The Veerman and Veldhuis-Diermanse (2001) scheme was used to distinguish task related messages, and the scheme by Fahy et al. (2000) was applied only to the task related messages.

Three independent researchers were used as coders. The coders were trained, but the details of the training protocol were not reported. The coders negotiated the final score, but the details of the negotiation protocol were not reported. Inter-rater reliability was assessed for initial scores (before negotiation) using Cronbach's alpha; however, the details of inter-rater reliability were not reported for separate topics or by coding scheme. Rather, the article merely mentions that inter-rater reliability ranged from 0.55 to 0.99 (presumably, across different topics and different coding schemes).

Collaborative learning messages dominated the discussion. In terms of the scheme by Veerman and Veldhuis-Diermanse (2001), the collaborative learning content was dominated by *evaluation*, *new theory*, and *new experience*—thus, both low and high levels of knowledge construction were well represented. In terms of the scheme by Fahy et al. (2000), the collaborative learning content was dominated by *reflections*, *statements*, and *horizontal questioning* (as one would expect from the correspondence indicated in Figure 4). Values for correlations between the results for the two schemes were not reported.

As in the earlier study by the same authors (Schellens and Valcke, 2005), a number of hypotheses relating to patterns of knowledge construction activities were tested. Once

again, the hypotheses were not based on a strong theoretical foundation, and, in my opinion, were primarily a form of presenting descriptive results.

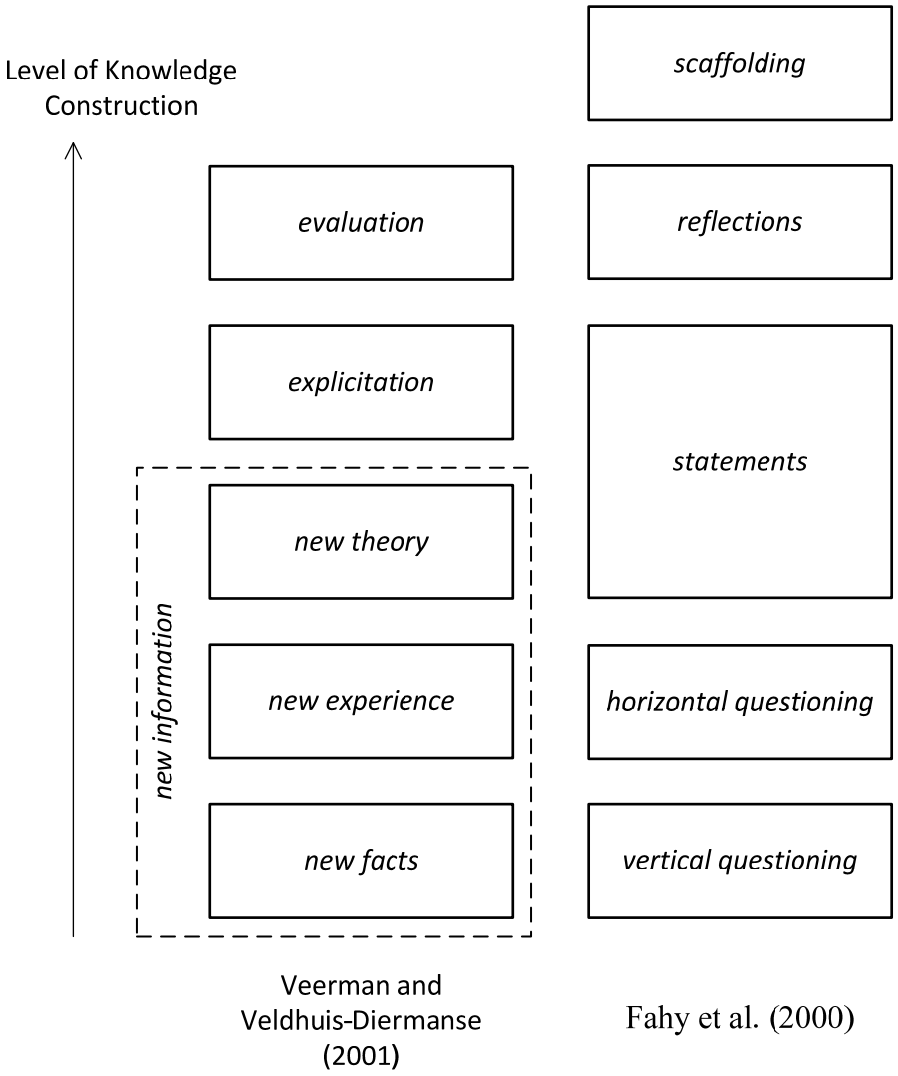

Figure 4. Mapping between Veerman and Veldhuis-Diermanse (2001) and Fahy et al. (2000) schemes.

In terms of the results of hypothesis testing, the results obtained by the two different coding schemes were consistent. As to the consistency possibly being due to using the same coders, the same criticism as that voiced at the end of section 2.6.3.2 applies.

2.6.4 Garrison (1991)—Cycle of Learning

Garrison (1991) proposed a model of knowledge construction that he called a "cycle of learning" (see Figure 5). The model suggests five phases of knowledge construction: problem identification (Phase 1), problem definition (Phase 2), exploration (Phase 3), establishing applicability (Phase 4), and integration (Phase 5).

Problem identification (Phase 1) refers to the initial awareness of a problem appearing in the discussion. In this phase, the contributors discover a dissonance between their knowledge and the issue at hand.

Problem definition (Phase 2) refers to the recognition of the problem by appraising and understanding the nature of the problem. In this phase, the contributors address the dissonance by establishing a shared problem definition.

Exploration (Phase 3) refers to the exploration of ideas related to the problem. In this phase, the contributors explore alternative approaches to resolving the problem (this involves critical analysis and integration of ideas).

Applicability (Phase 4) refers to assessing the applicability the alternatives identified in Phase 3 in terms of their applicability to solving the problem. The alterative selected can be seen as a hypothesis.

Integration (Phase 5) refers to testing the applicability of the solution selected in Phase 4 in solving the problem in practice. Thus, the hypothesis formulated in Phase 4 is tested. The outcome of this phase can be a successful resolution of the problem (thus, the hypothesis is confirmed). Alternatively (if the solution does not work), it triggers a new round of the cycle, starting from problem identification based on the new understanding attained in the round just completed.

Based on the Garrison's (1991) model of knowledge construction, Garrison, Anderson, and Archer (1999) developed an instrument for analysing the process of learning in CMC. Garrison et al. used three dimensions to describe the process of learning: cognitive presence, social presence, and teaching presence. The Garrison et al.'s scales for social presence and teaching presence are not discussed here as they are not relevant to knowledge construction. Garrison et al. used the term "cognitive presence" as a

synonym for critical thinking—thus; it was the dimension most relevant to knowledge construction.

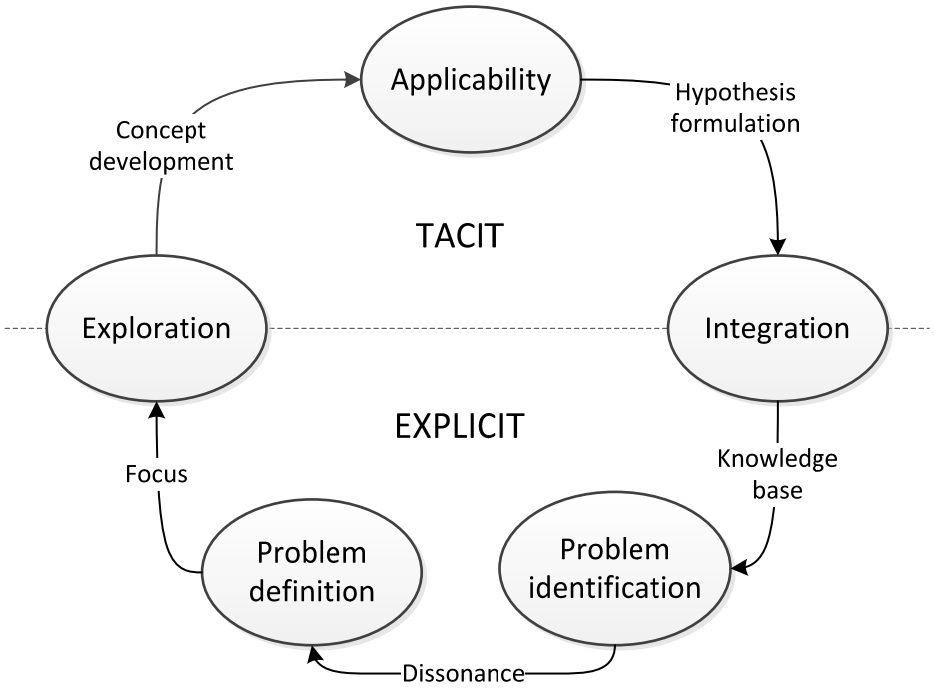

Figure 5. Cycle of learning model of knowledge construction by Garrison, based on Garrison (1991, p. 293).

The scale is summarized in Table 2-11. The levels of the scale correspond to the phases of knowledge construction in Garrison (1991) (*triggering event* corresponds to problem identification and problem definition, *exploration*—to exploration, *integration*—to applicability, and *resolution*—to integration).

Table 2-11 *Garrison et al.'s (1999) Dimension of Cognitive Presence*

Dimension	Scale	Levels
Cognitive Presence	Ordinal	*triggering event, exploration, integration, resolution*

Garrison et al. (1999) did not test the scale empirically. The known uses of the models developed by Garrison (1991) and by Garrison et al. (1999) are discussed in the following subsections.

2.6.4.1 Newman, Johnson, Webb, and Cochrane (1997)—Learning Quality in CSCL

The models developed by Henri (1992) and Garrison (1991) were used by Newman, Johnson, Webb, and Cochrane (1997) to develop a scale measuring the quality of learning. They applied the scale to analyse the quality of learning in a CSCL environment.

Table 2-12 The Structure of the Newman, Johnson, Webb, and Cochrane (1997) Model of the Quality of Learning

Dimension[a]	Scale	Levels	Comment
		New information[b]	
NI	Ordinal	-1,1	Contribution of new ideas
NS	Ordinal	-1,1	Contribution of new solutions
		Ambiguities[b]	
AC	Ordinal	-1,1	Clear statement
A	Ordinal	-1,1	Discuss ambiguities to clear them up
		Linking ideas / Interpretation[b]	
L	Ordinal	-1,1	Generating new meaning from the information collected
		Justification[b]	
JP	Ordinal	-1,1	Providing proof or examples
JS	Ordinal	-1,1	Justifying solutions or judgments
		Critical assessment[b]	
C	Ordinal	-1,1	Evaluation of a solution

Note: To save space, only a subset of the coding scheme is included in the table. Along with the dimensions at transcript level listed in the table, the coding scheme included Relevance, Importance, Bringing Outside Knowledge / Experience to Bear on Problem, Practical Utility, and Width of Understanding. Moreover, for the dimensions at transcript level included in the table only a subset of the corresponding dimensions at statement level is included. For details, refer to the original article.

[a]Dimensions at statement level. [b]Dimensions at transcript level.

The scale is summarized in Table 2-12. To measure the dimensions of the quality of learning for a transcript, sentences were coded and the resulting codes for the sentences in the transcript were added up.

Newman et al. (1997) analysed two transcripts: a transcript composed of messages generated over one semester of computer conferencing by undergraduate students in an Information Society course at a university in the UK and a transcript generated from recording face-to-face seminar in the same course. The unit of analysis was a transcript; the two transcripts were compared in terms of the dimensions of the quality of learning.

The computer conference transcript scored higher than the face-to-face seminar in Relevance, Importance, Bringing Outside Knowledge/Experience to Bear on a Problem, Linking Ideas/Interpretation, Justification, and Critical Assessment. The face-to-face transcript scored higher in Providing New Information. Thus, the computer conferencing environment offered better medium for knowledge construction.

Inter-rater reliability was not reported.

2.6.4.2 Garrison et al. (2001)—Critical Thinking and Cognitive Presence

Garrison, Anderson, and Archer (2001) adapted the Garrison et al. (1999) coding scheme to investigate cognitive presence in CMC (Garrison et al., 2001, used the term *cognitive presence* as a synonym for critical thinking). Garrison et al. (2001) hypothesized that cognitive presence (and hence, critical thinking) is affected by teaching and social presence. The Garrison et al.'s (2001) scales for social presence and teaching presence are not discussed here as they are not relevant to measuring knowledge construction.

Garrison et al. (2001) analysed three transcripts (see Table 2-13) comprising messages collected over 13 weeks of discussions by 26 students in a graduate program at a university in Canada. A message was used as a unit of analysis.

Table 2-13 *Transcripts Used in Garrison et al. (2001) Study*

	Contributors	Number of analysed messages
Transcript 1	13 graduate students from a workplace learning course	51
Transcript 2	6 graduate students from a health promotions course	20
Transcript 3	4 graduate students from a health promotions course	24

Inter-rater reliability was reported using Holsti's coefficient of reliability (CR) and Cohen's kappa (see Table 2-14). The background of the coders and coder training were not reported.

Table 2-14 *Inter-Rater Reliability in Garrison et al. (2001) Study*

	Cognitive presence		
Measured by	Transcript 1	Transcript 2	Transcript 3
CR	0.45	0.65	0.84
Cohen's kappa	0.35	0.49	0.74

The highest numbers of messages were in exploration (phase 2) and integration (phase 3) categories. Garrison et al. (2001) interpreted this by suggesting that because of the way the courses were conducted, the contributors were motivated to explore and understand the problems addressed in the discussions more than to solve them.

2.6.4.3 Schrire (2004)—Interaction and Cognition

Schrire (2004) investigated knowledge construction in term of interaction and cognition. Three models from cognitive theory were used by Schrire to assess the cognition process in computer conferences. The models were Bloom's Taxonomy (Furst, 1981), the Structure of Observed Learning Outcomes (SOLO) taxonomy (Chan, Tsui, Chan, & Hong, 2002), and the Garrison's (1999) Cycle of Learning. The scales are summarized in Table 2-15.

Table 2-15 *Knowledge Construction Scales Used by Schrire (2004)*

Coding scheme	Scale	Levels
Bloom's Taxonomy	Ordinal	*knowledge, comprehension, application, analysis, synthesis, evaluation*
SOLO Taxonomy	Ordinal	*prestructural, unistructural, multistructural, relational/extended abstract*
Cycle of Learning[a]	Ordinal	*initiation, exploration, integration, resolution*

[a]This scale was also introduced in section 2.6.4.

Schrire (2004) analysed 11 discussion threads (containing 87 messages) randomly selected from CMC discussions conducted over 20 months by graduate students enrolled in doctoral degree courses at a university in the USA. A message was used as a unit of analysis.

The inter-rater reliability was calculated using 40% of the messages analysed (thus, the reliability coder rated only 40% of the messages). Hoslti's coefficient or reliability (CR) was used. The inter-rater reliability was in the range of 0.55 – 0.74; the details for separate dimensions were not reported. Coder background and the details of coder training were not reported.

For all of the coding schemes involved, about half of the messages were rated at higher levels of knowledge construction. The same coder rated all messages, using the three coding schemes. Thus, the results may have reflected the coder's opinion regarding how the coding schemes related to each other.

2.6.5 Other Studies

This section discusses measures of knowledge construction that were not validated by being applied in multiple studies.

2.6.5.1 Järvelä and Häkkinen (2002)—Quality of Discussion

Järvelä and Häkkinen (2002) introduced a coding scheme to describe the type of communication and the level of knowledge construction in asynchronous CMC. The coding scheme is outlined in Table 2-16.

Table 2-16 *Coding Scheme by Järvelä and Häkkinen (2002)*

Dimension	Scale	Levels	Categories
Level of Discussion	Ordinal	*low, progressive, high*	
Type of Communication	Categorical		*comment, experience, new point/question, suggestion, theory*

The levels for the Level of Discussion dimension were not defined very clearly in the article, making it difficult to re-use the coding scheme or to interpret the results.

Järvelä and Häkkinen (2002) analysed 25 discussion threads (containing 342 messages) collected over 2 months of discussions by 70 pre-service teachers (40 accessing the system from the US and 30 from Finland) who used asynchronous discussions as part of their teacher training. A discussion thread was used as a unit of analysis.

Järvelä and Häkkinen (2002) found that the most common Type of Communication was *comment*, followed by *experience* report and *new point/question*. For the Level of Discussion dimension, lower level and progressive discussions were approximately equally common; high-level discussions were relatively rare.

The initial inter-rater reliability was 80% (the type of the measure was not explicitly reported, but most likely it was percent agreement). Coders then negotiated until agreement was reached.

The process used to recruit the coders was not reported. Presumably, the coders were the authors themselves. The details of coder training or coder negotiation protocol were not discussed.

2.6.5.2 Puntambekar (2006)—Analyzing Collaborative Interaction

Puntambekar (2006) introduced coding schemes for divergence of ideas, shared understanding and knowledge construction to investigate the process of collaborative learning. The scale for divergence of ideas is not discussed here as it is not relevant to knowledge construction. The coding scheme for shared understanding and knowledge construction is outlined in Table 2-17.

The Shared Understanding dimension referred to how the contributor's message related to messages from other participants. The Knowledge Construction dimension referred to learning occurring in the discussion.

Puntambekar (2006) analysed seven discussion threads (averaging 134 messages per thread) generated in asynchronous online discussions by 24 students enrolled in an

educational science course at a university in the USA. A message was used as a unit analysis.

Table 2-17 Coding Scheme for Shared Understanding (Puntambekar, 2006)

Dimension	Scale	Levels
Shared Understanding	Ordinal	*no thought, limited thoughtfulness, shows thoughtfulness, shows thoughtfulness supported by theory, extremely thoughtful opinions supported by theory*
Knowledge Construction	Ordinal	*no internalization; shows depth of understanding; shows depth of understanding supported by theory, with examples of application*

Note. Puntambekar (2006) used numbers for coding; to present their coding scheme compactly I replaced numbers by descriptive labels.

Puntambekar (2006) found that both shared understanding and knowledge construction tended to increase as discussions proceeded.

The inter-rater reliability for the Shared Understanding dimension ranged from 0.92 to 0.95 for different discussion threads (but the approach used to estimate inter-rater reliability was not reported). The inter-rater reliability for the Knowledge Construction dimension was not reported.

The process used to recruit the coders was not reported. Presumably, the coders were the authors themselves. The details of coder training were not discussed.

2.6.6 Summary and Comparison of the Measures of Knowledge Construction

The studies involving measures of knowledge construction reviewed in sections 2.6.1 to 2.6.5 were primarily descriptive studies. Only the studies by Schellens and Valcke (2005) and Schellens and Valcke (2006) involved hypotheses testing, but the hypotheses were not based on a strong theoretical foundation, and hypotheses testing was primarily used as a form of presenting descriptive results.

In this section I compare the measures of knowledge construction: first, from the perspective of their content, and then, from the perspective of replicability and reliability.

2.6.6.1 Measure Content

The content of the measures of knowledge construction introduced in sections 2.6.1 to 2.6.5 is compared in Table 2-18. I constructed the table by labelling and comparing categories and levels employed in the coding schemes related to knowledge construction, in the spirit of the constant comparative approach to qualitative data analysis (Corbin & Strauss, 2008).

Groups of categories and levels similar in meaning are presented in the "categories and levels" column and named in the "representative category" column (to name a group, I used the name of a category or level I considered to capture the best the common meaning of the whole group).

Only some of the dimensions in the coding schemes relating to knowledge construction used ordinal scales, thus suggesting the existence of levels of knowledge construction. The ordinal dimensions of knowledge construction were Cognitive (Henri, 1992), Critical Thinking (Newman et al., 1997 and Bullen, 1998), Phases of Knowledge Construction (Gunawardena et al., 1997), Cognitive Presence (Garrison et al., 2001), and Level of Discussion (Järvelä & Häkkinen, 2002). A number of the dimensions used scales that were categorical: Metacognitive Knowledge and Metacognitive skill (Henri, 1992); Focus and Constructive Activities (the dimensions by Veerman and Veldhuis-Diermanse, 2001); Type of Communication (Järvelä & Häkkinen, 2002); Knowledge, Metacognition, Conflict, Questioning, and Responses (the five dimensions by Hmelo-Silver, 2003); Knowledge Construction Categories (Pena-Shaff & Nichols, 2004). One may argue that the Collaborative Learning dimension of Veerman and Veldhuis-Diermanse (2001) was initially misclassified: it was initially treated as categorical, but the later study by Schellens and Valcke (2005) treated it as ordinal (and, indeed, one can clearly see that the categories suggest different levels of knowledge construction).

As to groups in Table 2-18, the group *management* relates to activities in support of knowledge construction rather than to knowledge construction per se. If the group

management is excluded, the remaining groups appear to reflect levels of knowledge construction, from *triggering event* at the lowest level to *resolution* at the highest level. The groups can be arranged as levels of a dimension with an ordinal scale, even though some of them include categories from dimensions measured on categorical scales.

Table 2-18 *Synthesis of Measures of Knowledge Construction*

Representative category	Categories and levels	Henri (1992), Bullen (1998), Hara et al (2000)	Pena-Shaff and Nichols (2004)	Gunawardena et al. (1997), Kanuka and Anderson (2007)	Veerman et al. (1999), Veerman and Veldhuis-Diermanse (2001)	Schellens and Valcke (2005)	Schellensand Valcke(2006)	Garrison et al. (1999), Garrison et al. (2001), Schrire (2004)	Newman et al. (1997)	Puntambekar (2006)
Triggering Event	*initiation, new information, question, sharing/comparing information, triggering event*		x	x	x	x	x	x	x	
Explicitation	*assertion, conflict, elementary clarification and in depth clarification, explicitation, exploration, exploration of dissonance among ideas, explanation*	x	x	x	x	x	x	x	x	
Integration	*consensus building, inference, interpretation, negotiation of meaning, shows depth of understanding, shows thoughtfulness and integration*	x	x	x		x		x	x	x
Evaluation	*evaluation, justification, judgment, reflection, strategies, testing and modifying proposed synthesis*	x	x	x	x	x	x		x	
Resolution	*achieving agreement, critical assessment, resolution,*			x		x		x	x	
Management	*assessment of the task, learner self-perception, means to succeed, planning, regulation, self-awareness*	x								

As seen from the table, only the coding scheme of Henri (1992) included categories falling into the *management* group. This was the consequence of treating the Henri's Metacognitive dimension as related to knowledge construction. The *evaluation* category from the Metacognitive dimension was placed under a group other than management—under *evaluation*. This was consistent with the opinion stated in the literature (Hara et al., 2000) that there is an overlap between the Henri's Cognitive and Metacognitive dimensions.

The majority of the coding schemes covered either all of the groups other than management or all of the groups minus one, suggesting an overall consistency between the dimensions.

2.6.6.2 Replicability and Reliability

Table 2-19 summarizes the studies from a methodology perspective. Only the studies that reported inter-rater reliability for dimensions measuring knowledge construction are included in the table. Of the studies not included in the table, Henri (1992) and Garrison (1991) only formulated coding schemes, and did not use them to code empirical data. Bullen (1998), Pena-Shaff and Nicholls (2004), Gunawardena et al. (1997), Kanuka and Anderson (2007), Veerman and Veldhuis-Diermanse (2001), and Newman et al. (1997) did conduct content analysis, but did not report inter-rater reliability or details relating to coder selection or training (presumably, the authors themselves acted as coders).

Of the studies that did report inter-rater reliability, Puntambekar (2006), Hara et al. (2000), and Veerman et al. (1999) reported reliability values for dimensions relating to knowledge construction that are acceptable or very close to acceptable according to methodological literature. The remaining four studies reported values that at least for some of the conditions or dimensions did not fulfil the criteria.

Krippendorff (2004) suggested that to ensure replicability content analysis studies should report coder background and use coders from common backgrounds (so that researchers attempting to replicate the study could use similar coders). Moreover, the details of coder training (if any) should be reported, so that the training can be

replicated. Coders should not be able to communicate during coding, because details of such communication cannot be controlled or replicated.

Table 2-19 Inter-Rater Reliability in Studies of Knowledge Construction

		Coder Training		Inter-rater reliability			Negotiation	
Study	Coders	Conduct-ed	Details of protocol	Conduct-ed	Method[a]	Range	Conduc t-ed	Details of protocol
Hara et al. (2000)	Three coders, background not reported	Not reported		Yes	Percent agreement	0.71-0.75	Yes	Not reported
Veerman et al. (1999)	Not reported	Not reported		Yes	Cohen's kappa	0.74	Not reported	
Schellens and Valcke (2005)	Three independent researchers	Yes	Not reported	Yes	Cronbach's alpha	0.45-0.93	Yes	Not reported
Schellens and Valcke (2006)	Three independent researchers	Yes	Not reported	Yes	Cronbach's alpha	0.55-0.99	Yes	Not reported
Garrison et al. (2001)	Two coders, background not reported	Not reported		Yes	CR	0.45-0.84	Not reported	
				Yes	Cohen's kappa	0.35-0.74		
Schrire (2004)	Two coders, background not reported	Not reported		Yes	CR	0.55-0.74	Yes	Not reported
Puntambekar (2006)	Two coders, background not reported	Not reported		Yes	Not reported	0.92-0.95	Yes	Not reported

Note. CR=Holsti's coefficient of reliability.

[a]Thresholds for acceptable inter-rater reliability were suggested as follows in the literature: percent agreement > 0.8 (Riffe et al., 2005), CR > 0.8 (Riffe et al., 2005), Cohen's kappa > 0.75 (Rourke & Anderson, 2001), and Cronbach's alpha > 0.7 (Nunnally & Bernstein, 1994).

None of the studies that reported acceptable (according to methodological literature) levels of inter-rater reliability reported the details of coder training (or even reported if such training was conducted or not). It is doubtful that high inter-rater reliability could be achieved without training; the details of coding schemes as reported clearly allowed multiple interpretation. The two studies that did report conducting coder training, Schellens and Valcke (2005) and Schellens and Valcke (2006), did not report the details of the training.

The studies that reported acceptable (according to methodological literature) levels of inter-rater reliability did not report coder selection procedures or background of the coders. In fact, it appears that authors themselves acted as coders; if that was the case, the authors' experiences in discussing and formulating the coding scheme constituted coder training. Clearly, this kind of coder training cannot be documented and cannot be replicated by others. Thus, the relatively high inter-rater reliability values reported in these studies have little value in terms of ensuring replicability.

If the coders were the authors themselves, it may be impossible to replicate the studies, because the meanings associated with interpreting the indicators were based on the understanding achieved in conceiving and designing the study and the coding scheme, rather than on just reading the coding scheme and on documented, replicable training. An understanding attained based solely on reading the coding scheme is likely to differ from the understanding by the researchers who designed the coding scheme.

Schellens and Valcke (2005) and Schellens and Valcke (2006) were the only studies that reported coder background, and were the only studies in which it was clear that the coders were not the authors. The description of the coder background was limited to stating that the coders were "independent researchers". Clearly this description is not sufficient to allow replication—the background of the "independent researchers" was not reported. In both studies, inter-rater reliability values were reported as ranges; inter-rater reliabilities for individual dimensions were not reported. It is not clear if the range included inter-rater reliability for separating task-related messages from messages that are not task related (e.g., social exchanges) and thus not appropriate for coding on knowledge construction. Distinguishing task-related messages from non-task-related messages does not involve the interpretation of complex concepts and thus can be done much more reliably than coding for levels or categories of knowledge construction. Overall, only the studies by Henri (1992) and Veerman and Veldhuis-Diermanse (2001) explicitly addressed the issue of separating task-related content from non-task related content, although it was relevant for all of the studies.

Schellens and Valcke (2005) and Schellens and Valcke (2006) compared different coding schemes for knowledge construction and found them to be consistent. However, the same coders were used for the different coding schemes that were compared. Thus,

common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003) was likely, because the results may have reflected the opinion of the coders regarding how the coding schemes were related. As learning to use a new coding scheme requires considerable cognitive effort, it would be natural for the coders (consciously or unconsciously) to continue using the rating scheme they already knew, and to map the results to the new coding scheme, resulting in artificially high consistency.

2.7 Knowledge Construction in KM

The dichotomy between knowledge and data is central in KM, and the ascension from data (facts out of context), to information (facts making sense in context), and to knowledge (capacity to act intelligently) is seen as knowledge creation (Dalkir, 2011). A similar hierarchy considered in KM is the ascension from *know what* (making sense of the situation), to *know how* (capacity to act), to *know why* (capacity to act intelligently), and to *care why* (capacity to innovate). The ascension from data to knowledge (and the ascension from *know what* to *know why* and *care why*) in context of an organization improves the ability of the organization to create value for its customers. Knowledge creation occurs continuously. The organization enhances its ability to create value by facilitating and promoting the ongoing knowledge creation (Davidson & Voss, 2002; Hwang, 2003).

The concept of knowledge creation in knowledge management is clearly distinct from the concept of knowledge construction in e-learning. Knowledge creation in KM emphasizes knowledge retention and build up at an organizational level; knowledge construction in e-learning emphasizes knowledge acquisition by individuals and by groups of interacting individuals. Knowledge creation in KM includes activities such as extracting value from organizational data (and converting it into information) by using data mining (Dalkir, 2011)—something that is clearly not within the scope of knowledge construction as it is understood in e-learning. Conversely, communities of practice in KM are similar to groups of learners using CMC to construct knowledge. Indeed, social constructivism is frequently mentioned in KM literature in context of discussing the use of communities of practice (Dalkir, 2011).

Knowledge creation in KM ultimately results in individuals acquiring knowledge (that is, learning), and thus acquiring the capacity to act intelligently. Such learning occurs in social context; therefore, there is a clear parallelism between learning as part of KM at organizations and social constructivist learning in e-learning. Exploring this parallelism in depth is outside the scope of this study; rather, the rest of this section is limited to reviewing KM literature that is particularly relevant to this study.

In the rest of this section, I discuss a study that introduced a model describing the role of knowledge construction in KM (McAdam & McCreedy, 1999), followed by a study that suggested strategies for enhancing knowledge construction in organizations (Hwang, 2003). The section concludes by discussing an empirical study of online discussions conducted from a KM perspective and covering both active participants and lurkers (Milne & Callahan, 2006).

2.7.1 McAdam and McCreedy Model of the Role of Knowledge Construction in KM

McAdam and McCreedy (1999) developed a model describing the role of knowledge construction in KM (see Figure 6). The model emphasizes the dichotomy between the scientific and the social paradigms of knowledge construction. The scientific paradigm views knowledge as a body of objective facts, and stems from the constructivist theory as it is used in philosophy of science (Phillips, 1995). The scientific paradigm suggests that knowledge is constructed by the global research community; from the perspective of an organization, such knowledge is imported, rather than constructed locally. For such imported knowledge, local knowledge construction is limited to making sense of the imported knowledge in the local context. In contract, the social paradigm views knowledge as constructed locally, by the learners themselves, in context of the organization. The social paradigm corresponds to the social constructivism view of learning introduced in section 2.5.2.

As seen in the model in Figure 6, McAdam and McCreedy (1999) suggested that both the scientific paradigm and the social paradigm are relevant to explaining knowledge construction at organizations. Moreover, knowledge dissemination (making knowledge available to the stakeholders), knowledge embodiment (capturing knowledge as

documents, artefacts, and procedures), and knowledge use (applying knowledge to promote organizational goals) both benefit from and contribute to knowledge construction. Activities involved in knowledge dissemination, embodiment, and use promote discussions resulting in social knowledge construction. At the same time, discussions involved in social knowledge construction transform scientific knowledge to make it relevant to the organization, and result in knowledge dissemination, embodiment, and use (Kautz & Thaysen, 2001; McAdam & Reid, 2001).

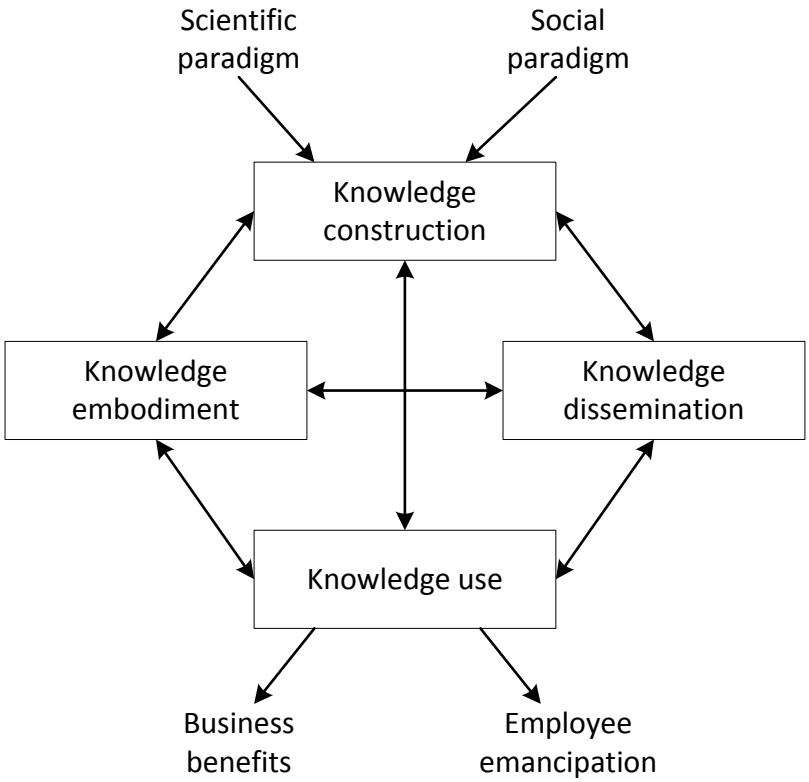

Figure 6. Knowledge construction in knowledge management—A model by McAdam and McCreedy (1999) (p. 103).

2.7.2 Enhancing Knowledge Construction at Organizations

Hwang (2003), based on the social constructivism theory, suggested six strategies to enhance knowledge construction at organizations.

First is reflection—the employees need to be aware of knowledge construction and aware of acquiring new knowledge and of decisions to use it in practice. Reflection enables the employees to continuously improve knowledge construction.

Second is promoting knowledge transfer and knowledge construction by providing opportunities and media for collaboration. This strategy is motivated by the social constructivism theories suggesting that social interactions enable knowledge construction (Kim & King, 2004).

Third is providing employees with opportunities to have a variety of relevant experiences, thus promoting the construction of tacit knowledge. The sharing of such tacit knowledge via social interactions (the second strategy) further enhances knowledge acquisition throughout the organization.

Fourth is helping the employees to develop and to share mental models. This is achieved by making the employees aware of relevant conceptual frameworks (such as the Porter's value chain model) that can be used as a basis for mental model building and by organizing workshops and activities offering employees opportunities to externalize their mental models (thus externalizing, testing, and sharing their knowledge) (Cabanero-Johnson & Berge, 2009).

Fifth, enable and promote communities of practice spanning organizational boundaries (Wenger, 1998). Communities of practice bring together employees interested in a particular topic from throughout the organization (or, possibly, from different organizations). This results in knowledge transfer via social knowledge construction. Moreover, it enables the exchange of tacit knowledge between employees that are otherwise unlikely to come in contact. One way to enable communities of practice is by providing media for collaboration (the second strategy).

Six, encourage system thinking. Encourage the employees to view problems in a broad context, in the context of the whole organization, rather than in contexts of particular narrow functions (Rowley, 2000; Senge, 2006).

2.7.3 An Empirical Study of Online Discussions Conducted From a KM Perspective

Milne and Callahan (2006) conducted a descriptive study of an online community of practice devoted to knowledge management in the public sector. About 180 of the community members who responded filled in an online questionnaire; furthermore,

nine of the members who frequently contributed to the discussions were interviewed over the phone.

Most of the members were lurkers—only about 22% of the respondents regularly contributed to online discussions. Yet, the community affected the practice—over 21% of the respondents indicated that the discussions triggered KM initiatives at their organization (however, a breakdown of this number between lurkers and active participants was not given). The respondents indicated that they learned both from the postings they read and from interactions at private networks formed by members that met online via the community. The respondents found that the main value of the participation was the networking opportunities the community provided and the ability to learn from the experience of others.

Frequent contributors were the most motivated to participate in discussions when the issues discussed had immediate relevance to on-the-job problems they faced and when the discussion pushed the boundaries of the accepted practice. Discussions that were academic in nature and discussions involving the participants adopting inflexible "black and white" positions were regarded negatively. Frequent contributors responded rapidly to posts of interests and only rarely took time to reflect.

In my opinion, even though Milne and Callahan (2006) did not refer to social constructivism in the article, the features of the discussions such as spontaneous responses, focus on practice, learning from the experience of others, and the formation of off-line networks further contributing to learning were consistent with social constructivist learning.

2.8 Perceived Information Quality and Information Integrity

In my view, interactions in social constructivist learning can be seen as exchanges of information. This is particularly clear in social constructivist learning using text based CMC—transcripts of CMC discussions are readily available and can be treated as information. From an information perspective, learning in social constructivist learning happens as learners externalize their knowledge as information and internalize the information externalized by other participants.

Further, I take the view that in a CMC discussions in context of constructivist learning, messages of sharing, explanation, negation, evaluation, and agreement form an information flow in the form of a Socratic dialogue. The direct participants in the interaction are not the only consumers of this information; rather, usually there are many lurkers, who access the information flow as it develops or access it after the discussion is over (essentially, as a document). As demonstrated in section 2.7.3, lurkers might considerably outnumber the active participants. Thus, an important group of stakeholders, lurkers (or "legitimate peripheral participants"—see section 2.4), learn from a CMC discussion solely by receiving the information flow generated by the discussion as an input.

Not all information is the same, but information may be of different quality (Fisher, Lauría, Chengalur-Smith, & Wang, 2012). The quality of learning (particularly, by lurkers) in a CMC discussion can then be related to the quality of information generated in the discussion, and the quality of information can be related to the knowledge construction in the discussion (this aspect is discussed in more detail in sections 1.6 and 3.3.1).

Information quality can be considered from two perspectives: objective quality and perceived quality. On the one hand, one can consider the data content of the information artefact, such as a document. From the perspective of data content, the quality of the information artefact is determined by the usefulness of the content in decision making once the content is understood by the user: whether the data accurately reflect the facts in the real world, whether the data reflect the facts at the targeted moment in time, whether there is enough data, whether different data items are consistent with each other and so on. The data content quality perspective can be seen as objective; information quality from that perspective does not depend on the perceptions of the artefact by the users of the artefact. On the other hand, from the perspective of perceived information quality, the quality of the information artefact is determined by user perceptions. An information artefact perceived by its users as relevant, understandable, believable, useful, and so on is seen as being of high quality. The dichotomy between the data content quality perspective and the perceived information quality perspective outlined above is based on the discussion of perceived information quality by Nicolaou and McKnight (2006). Nicolaou and McKnight

distinguished “information integrity”, “data quality”, and “information quality”; however, the distinction between “data quality” and “information quality” was not clear enough, as, for example, relevance was included under both; thus, I simplified their classification to a dichotomy, with data content quality perspective corresponding to “information integrity”, and perceived information quality perspective—to a combination of “data quality” and “information quality”.

For assessing the quality of the information flow generated in online discussions, both perspectives are relevant. When facts are mentioned in a discussion, if wrong facts are not detected as such, the quality of the discussion is clearly undermined. On the other hand, if the members (both frequent contributors and lurkers) perceive the information they obtain from the discussion as irrelevant or not believable, they are less likely to internalize the information as knowledge; thus, perceived information quality is also relevant.

For health-related discussions, it is essential that health-related interventions and behaviours suggested in the discussions actually lead to better health outcomes; this aspect is discussed in section 2.9. The rest of this section focuses on perceived information quality.

McKinney, Yoon, and Zahedi (2002) suggested a conceptualization and a measure of perceived information quality of web content. Even though an alternative measure of information quality was proposed by Lee, Strong, Kahn, and Wang (2002), it was more suitable for database reports than for less structured web content, and focused on aspects, such as ease of manipulation, clearly not relevant to online discussions. Therefore, here I focus on the conceptualization and measure of perceived information quality by McKinney et al.

McKinney et al. (2002) developed a multidimensional conceptualization and measure of perceived information quality. The initial pool of items was formulated based on reviewing the literature. Several rounds of pilot tests followed by factor analysis were conducted to purify and to test the instrument, resulting in six dimensions of perceived information quality and the corresponding measurement items. McKinney et al. proceeded to further simplify the conceptualization and the measure, at the cost of

losing relevant content, because they ultimately needed to use the resulting dimensions in a second order SEM model. This consideration did not apply to the present study; therefore, I used the outcome of the first phase of the McKinney et al. measure development, comprising six dimensions: relevance, understandability, reliability, adequacy, scope, and usefulness, with the items tested in three pilot studies.

In the following subsections I briefly discuss the McKinney et al. (2002) six dimensions of perceived information quality.

Relevance. McKinney et al. (2002) defined relevance as "concerned with such issues as relevancy, clearness, and goodness of the information" (p. 301). According to WordNet (a lexical database of English maintained by Princeton University in the US, http://wordnetweb.princeton.edu/perl/webwn) and according to Merriam-Webster (Relevant, n.d.) and Chambers (Brookes, 2003) dictionaries, the difference in meaning between relevance and relevancy is unclear, and the two appear to be synonyms expressing largely the same meaning. A Google search conducted at the time of writing (March 2012) suggested that the term "relevance" is about ten times more common than "relevancy". Thus, the definition by McKinney et al. appears to be circular; moreover, the term was defined via its less commonly used synonym.

For the purposes of this study, I follow the definition of relevance given by the Chambers dictionary (Brookes, 2003), which is "bearing upon or applying to the matter in hand, pertinent", and further define relevance as the information applicable to resolving the problem faced by its user. This definition is consistent with the items used in the McKinney et al. (2002) measure, suggesting that relevant information is applicable, related, and pertinent to the problem faced by the user. Moreover, this definition is consistent with the one suggested by Knight and Burn (2005), who defined relevance as the "extent to which information is applicable and helpful for the task at hand" (p. 162).

Understandability. McKinney et al. (2002) provided no initial definition for understandability, because the understandability dimension was added based on the analysis of pilot data. Knight and Burn (2005) defined understandability as the "extent to which data are clear without ambiguity and easily comprehended" (p. 162). The

items used in the McKinney et al. measure of understandability suggested that understandable information is clear in meaning, easy to comprehend, and easy to read (also, see the discussion in section 3.2.2).

Reliability. McKinney et al. (2002) defined reliability as "concerned with the degree of accuracy, dependability, and consistency of the information" (p. 301). The items used in the McKinney et al. measure of reliability suggested that reliable information is trustworthy, accurate, and credible; however, "consistency" was not represented in the measure directly. (I believe that this is appropriate, because consistency is more relevant to structured information than to mainly unstructured information on the Internet or in online discussions.)

Adequacy. McKinney et al. (2002) provided no initial definition for adequacy, because the adequacy dimension was added based on the analysis of pilot data. Lin, Lu, Wang, and Wei (2011), in a related study, defined adequacy as "concerned with the completeness of information" (p. 618). The items used in the McKinney et al. measure of adequacy suggested that adequate information is sufficient, complete, and covers the necessary topics.

Scope. McKinney et al. (2002) defined scope as "scope evaluates the extent of information, range of information and level of detail provided by the website" (p. 301). The items used in the McKinney et al. measure of scope suggested that information with scope covers a wide range, contains a wide variety of topics, and covers a number of different subjects.

Usefulness. McKinney et al. (2002) defined usefulness (in the context of e-commerce) as "users' assessment of the likelihood that the information will enhance their purchasing decision" (p. 301). The items used in the McKinney et al. measure of usefulness suggested that useful information is informative to the purchase decision and valuable to the purchase decision. For the purposes of this study, I reworded the definition by McKinney et al. for use in a more general context, and defined usefulness as users' assessment of the likelihood that the information will enhance their ability to address their concerns.

2.9 Information Integrity and Evidence-Based Medicine

Because not all online communities are supervised by health practitioners, health-related interventions and behaviours suggested by the discussions do not necessarily reflect the current state of scientific medical knowledge. This can be seen as an information integrity problem—recommendations that are in agreement and are ultimately based on the current state of scientific medical knowledge can be described as information of high integrity. Other recommendations can be seen as information lacking integrity or as information of uncertain integrity.

In healthcare, recommendations known to be based on the current state of scientific medical knowledge are described as evidence-based medicine (Rosenberg & Donald, 1995). Recommendations that are not based on the current state of scientific medical knowledge are seen as anecdotal; physicians are expected to base their practice on evidence-based medicine.

To establish recommendations that can be described as evidence-based medicine, systematic reviews of medical research literature are conducted on an ongoing basis, and the results are published as databases. Describing and comparing all of the existing methodologies for such systematic reviews is beyond the scope of this thesis; instead, I briefly describe the methodology developed by Cochrane Collaboration (The Cochrane Collaboration, 2013)—an influential international organization promoting systematic reviews of medical research literature. In New Zealand, the database of systematic reviews maintained by the Cochrane Collaboration (http://www.thecochranelibrary.com) was made available to all New Zealanders free of charge in 2006 (Jordan, Jeffery, & Farquhar, 2006) as a result of a national license purchased by the New Zealand Government. (The database remained available at the time of writing, March 2012.)

Cochrane Collaboration publishes the "Cochrane Handbook for Systematic Reviews of Interventions" detailing the methodology (Higgins & Green, 2008), henceforth in this section referred to as the Handbook. The role of the authors of systematic reviews is, in some respects, similar to the role of coders in content analysis. The issue of who could be a suitable author is not addressed in the Handbook. Nonetheless, all the reviews I

could access in the Cochrane Collaboration database were authored by individuals affiliated with medical institutions, even though the editorial procedures (Higgins & Green, 2008, p. 11) do not explicitly impose any restrictions on who can be a reviewer.

The Handbook does suggest that prospective review authors undergo training, and at the time of writing the Cochrane Collaboration web site provided extensive information on training opportunities.

The unit of analysis in a systematic review is a research publication (a study). The Handbook discusses in detail the approach to selecting the studies for analysis (a procedure corresponding to sampling in content analysis). The studies are selected based on the focus of the review (the questions the review intends to address), the study participants, the interventions, the outcomes for the participants, and the research design. The recommended sources of studies are primarily online bibliographic databases, although other sources (such as medical trial registers) are also suggested.

Once the studies are selected, specific information (to be analysed in the following stage) is extracted and presented in a uniform way, including information about the methods, the participants, the interventions, and the outcomes. This step is similar to extracting coding units from sampling units in content analysis.

Once the information is extracted and, using the vocabulary of content analysis, coding units are obtained, they are assessed for bias (a stage similar to coding in content analysis), such as selection, performance, attrition, detection, and reporting bias (I do not describe what these biases constitute, because that would require me to introduce the details of experimental research methodology, which is clearly beyond the scope of this literature review).

Similarly to content analysis, the Cochrane methodology employs inter-rater reliability as evidence of the validity of the finding. Nonetheless, comparing to the approach to content analysis by Krippendorff (2004), there is much less emphasis on inter-rater reliability. Reviews that do not involve the assessment of inter-rater reliability (e.g., with a single author) are allowed. Inter-rater reliability is assessed only for the most important data (such as the key outcomes), and insufficient inter-rater reliability does not preclude the review from being published. Although Krippendorff suggests that in

case of problems with inter-rater reliability the coding scheme should be updated and the data should be recoded using different coders, the Cochrane methodology emphasizes carefully recording any disagreements. The Cochrane methodology does not preclude the same individuals who designed the study from assessing the studies reviewed for bias (the activity corresponding to coding in content analysis). Thus, any inter-rater reliability values obtained in a Cochrane review are likely to reflect the shared understanding obtained in designing the study.

The above comparison between the Cochrane methodology and the approach to content analysis suggested by Krippendorff (2004) is based on how the methodologies are formulated, not on the comparison of the actual practice. The actual practice may be closer than the stated methodologies. In particular, it is likely that Cochrane reviews with multiple authors and with inter-rater reliability assessed are easier to publish; moreover, as can be seen from the review of content analysis practice in studies of knowledge construction presented in section 2.6, content analysis studies do not always reach the inter-rater reliability benchmarks suggested by Krippendorff. It would be of interest to study and to compare the actual practice, but this is clearly out of the scope of the present study.

The last stage in the Cochrane methodology involves presenting a summary of results in tabular form, which is also a common way to present the results on content analysis (Krippendorff, 2004, p. 192). The Handbook suggests specific formats in which the results are to be summarized. The ultimate outcome of a review is a compact presentation of the studies included in the analysis, and, for each study, an assessment of possible sources of bias. Studies that are sufficiently unbiased are seen as presenting evidence-based results.

To ensure that recommendations suggested in online health support group discussions are evidence based, a moderator would need to assess them by evaluating them against the databases of medical knowledge such as the databases provided by Cochrane Collaboration. A less expensive option would be to have a medically qualified participant occasionally present, who would assess the integrity of the information based on her expertise. It remains an open question, though, if the regular participants themselves, in discussions conducted at high levels of knowledge construction, would

be able to challenge recommendations that are not evidence based and thus maintain the integrity of information in the discussion.

2.10 Knowledge Gap

Health support groups enable individuals to be involved to a greater extent in addressing their health-related issues. The trend towards the patient-centred care paradigm in healthcare highlights the role of peer-to-peer online health support groups (see section 2.3).

There is evidence that individuals involved in health support groups use them primarily to obtain information, rather than to seek moral support or for any other purposes (as discussed in section 2.3). There is also evidence that most of the users of online discussions are legitimate peripheral participants (or "lurkers"), and thus consume the discussion as an information flow, rather than actively contributing (see sections 2.4 and 2.7.3). Even though decisions based on the information gained from health support group discussions can have important consequences for individuals, the quality of information flow (e.g., in terms of perceived information quality and information integrity, see sections 2.8 and 2.9) generated by peer-to-peer online health support groups has never been studied.

Knowledge construction in online discussions supporting communities of practice has been highlighted as contributing to knowledge creation at organizations (see section 2.7.1). Even though there are multiple studies of knowledge construction in online discussions in educational settings that presented evidence of knowledge construction occurring (see section 2.6), studies of knowledge construction in online discussion supporting communities of practice are rare (I found only one study, see section 2.6.2.1, and the study had inconclusive results and methodologically was not robust). There are no studies of knowledge construction in health support group online discussion. More specifically, the effect of knowledge construction in health support group online discussion on the quality of information flow generated by such discussion has never been considered.

2.11 Health Literacy

Knowledge construction in health support group online discussions involves the contributors both writing and posting and reading and interpreting messages. Clearly, the ability of the contributors to write meaningful health-related messages and to interpret health-related messages by others depends on their literacy, in particular, on their health literacy. This section reviews the literature on health literacy, emphasizing the literature that is particularly relevant to understanding the relationships between health literacy and knowledge construction in health support group online discussions. The section concludes by further delineating the scope of the present study, to explain how the purpose and the high-level hypotheses of the present study (formulated in sections 1.2 and 1.6) relate to health literacy.

2.11.1 Conceptualization of Health Literacy

Health literacy is the ability of individuals to handle words and numbers in context of health-related communication. A range of definitions of health literacy was reviewed by Sorensen (2012). Some of the definitions of health literacy emphasize the ability to understand and to convey health-related information, such as the definition in the health promotion glossary by Nutbeam (1998): "The cognitive and social skills which determine the motivation and ability of individuals to gain access to understand and use information in ways which promote and maintain good health" (p. 357) (the health promotion glossary was also published as an official World Health Organization document, refer to http://www.who.int/healthpromotion/about/HPG/en/).

Other definitions imply the possession of health-related knowledge along with the ability to understand and to convey health-related information, such as the definition by Kickbusch, Wait, and Maag (2005) given in their position article presenting the views of Alliance for Health and the Future, "The ability to make sound health decision(s) in the context of everyday life—at home, in the community, at the workplace, the healthcare system, the market place and the political arena. It is a critical empowerment strategy to increase the people's control over their health, their ability to seek out information and their ability to take responsibility" (p. 8). A similar view was taken by Helitzer, Hollis, Cotner, and Oestreicher (2009), who conducted formal content

validation of written health information materials about cervical cancer and asserted that "Health literacy requires reading and writing skills as well as knowledge of health topics and health systems" (p. 1). Connor (2009), based on a survey of Internet users searching for health-related information, concluded that health literacy goes beyond reading ability, and includes communication ability and the knowledge of scientific vocabulary. The Mancuso's (2008) definition, which summarizes a conceptual model that she formulated by synthesizing the findings of a systematic review of literature on health literacy, stands out by emphasizing the process, rather than the ability: "A process that evolves over one's lifetime and encompasses the attributes of capacity, comprehension, and communication. The attributes of health literacy are integrated within and preceded by the skills, strategies, and abilities embedded within the competencies needed to attain health literacy" (p. 250).

Taking the broader view of health literacy as reflecting both the ability to acquire and to impart health-related knowledge and the possession of health-related knowledge, one may argue that knowledge construction in health support group online discussions not only depends on health literacy (the contributors should be able to contribute in writing and should understand each other's messages), but may also result in improved health literacy (by improving the contributors' and the lurkers' health-related knowledge). Moreover, if one adopts the Mancuso's view on health literacy, the knowledge construction process can be seen as part of the Mancuso's health literacy process, and , therefore, knowledge construction in health support online discussions is a part of health literacy, rather than its antecedent or its outcome. In the following discussion, though, I adopt the more conventional view of health literacy as an attribute, rather than as a process.

2.11.2 Health Support Online Discussions as an Antecedent of Health Literacy

The existing research on health literacy emphasizes the ability of individuals to interpret documents (off-line and online) with health-related information (see, for example, Helitzer et al., 2009 and Park, Rodgers, & Stemmle, 2013). The content of such documents is fixed once the documents are published. Often, such content is not tailored for different cultural groups, and it is not tailored to individuals.

The importance to tailor messages to the individuals was highlighted by Gustafson et al. (1999), based on testing a prototype adaptive decision support system with HIV-positive patients. Knowledge constructed in an online discussion involving an individual addresses the individual's specific comments and context (including cultural context); therefore, it is adaptive, it is tailored to the individual (in the sense of the "client-oriented" model of health-related online engagement proposed by Woerkum, 2003). Moreover, the resulting document (the transcript of the discussion) is likely to be suitable for individuals who do not take part, but are similar to the individuals who do (in other words, suitable for some of the lurkers).

Because of the adaptive nature of information provided via online discussions, health support group online discussions may promote health literacy by making acquiring health-related knowledge less demanding, responding to the Sorensen's et al. (2012) call to "improve the readability of the system" (p. 10). Baker (2006) explored the existing conceptualizations and measures of health literacy and formulated a conceptual model postulating the major antecedents and outcomes of print and oral health literacies. According to the model, the complexity of the printed and spoken messages in the healthcare environment negatively affects health literacy. Aguilera, Perez, and Alonso Palacio (2010), based on their study of readability levels of written diabetes health education materials, found that materials related to living with diabetes created to emphasize readability by a nonprofit organization were still above the reading level of the intended population. Helitzer et al. (2009), who conducted a health literacy assessment of written health information materials on cervical cancer prevention, reached a similar conclusion—the materials were too complex for the intended readers. The American Medical Association, based on a Delphi study of the consequences of poor health literacy in the United States (involving a panel of health literacy experts), emphasized the importance of improving communication with low-literacy patients (Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs of American Medical Association, 1999). Health support group online discussions could contribute to heeding this call by providing health-related information in a form matching the level of health literacy of the information consumers.

The importance of tailoring health information to the cultural group was highlighted by Kreuter, Lukwago, Bucholtz, Clark, and Sanders-Thompson (2003), based on testing

tailored cervical cancer awareness materials with African American females. Moreover, Guidry, Fagan, and Walker (1998), based on their analysis of cancer prevention materials content, and Betancourt, Green, Carillo, and Ananeh-Firempong (2003), based on their review of publications focusing on sociocultural barriers to care, emphasized that health-related materials should be culturally appropriate to be effective. In health-related online discussions, knowledge is negotiated and created by the contributors, and, therefore, is likely to be culturally appropriate to them. Further, McCray (2005), in her review of literature on healthcare information and health literacy, emphasized that health-related literacy is best promoted via culturally appropriate relevant content focused on actions and behaviours rather than underlying principles; it would be reasonable to expect that in online discussions the contributors are more likely to exchange experiences relating to actions and behaviours, than to lecture each other about the scientific principles behind the recommended health-related behaviours.

A number of authors have emphasized the role of social support in promoting health literacy. Paasche-Orlow and Wolf (2007), based on their review on medical and public health literature on health literacy, asserted that social support promotes health literacy; such social support could be gained in an online discussion. In a similar vein, Kickbush and Maag (2008), in their article in International Encyclopedia of Public Health, emphasized the role of community in promoting health literacy; the community could be an online community.

The role of interactive virtual environments in promoting health literacy has been explicitly highlighted by a number of studies. Norman et al. (2008), based on their survey of high school students, concluded that student engagement in virtual health instruction classrooms (which feature discussion forums) results in improved health literacy. Park et al. (2013) conducted a content analysis of tweets from health-related organizations and found large numbers of twitter messages related to health literacy at health-related organizations and community groups; Park et al. highlighted the use of simple language in such messages. November and Day (2012), based on their educational design research study of health and musicology students, suggested that the effectiveness of peer to peer online interactions in promoting students' health literacy is moderated by the students' digital literacy.

2.11.3 Health Literacy as an Antecedent of Knowledge Construction

A number of authors have expressed views suggesting that health literacy enables individuals to engage in meaningful discussions on health-related matters, a view that suggests that health literacy is an antecedent of knowledge construction in health support group online discussions. Zarcadoolas, Pleasant, and Greer (2005), based on content analysis of mass media publications devoted to the anthrax threat in the United States during 2001, concluded that a health literate person is able to apply health concepts and information to novel situations and to participated in public dialogues about health.

Nutbeam (2000), based on his review and analysis of educational programs and communication strategies intended to promote health literacy, distinguished interactive health literacy (ability to convey and derive health-related meanings in everyday communication) and critical health literacy (ability to critically analyse health-related information). Clearly, both interactive and critical health literacy are needed to fully realize the potential of health-related online discussions (unless some of the contributors possess these dimensions of health literacy, knowledge construction is not likely to happen). On a similar note, Manganello (2008), based on a review of literature related to health literacy of adolescents, emphasized the dimension of media literacy, which he described as the ability to critically evaluate media messages.

Porter and Edirippulige (2007), based on a survey to examine the patterns of Internet use by parents of deaf children seeking hearing-loss-related information, found that better educated (and thus, more literate) parents were more likely to find information on the Internet that had a major influence on how they managed their child's hearing loss. Chan, Matthews, and Kaufman (2009), based on analysing the content of a health-related web site, identified the levels of health literacy required to access different parts of the site (thus emphasizing that health literacy enables access to health-related online materials).

2.11.4 Health Literacy and the Purpose of the Present Study

The purpose and the high-level hypotheses of the present study (see sections 1.2 and 1.6) did not involve health literacy. Yet, as follows from the literature reviewed in

sections 2.11.2 and 2.11.3, health literacy may be related to knowledge construction, suggesting possible directions for future research. In particular, health literacy of discussion contributors both enables them to contribute and improves as the result of their contribution, as they build up their tacit knowledge by engaging in health-related discussions. Lurkers do not contribute, and thus their health literacy does not result in knowledge construction captured in the discussion content; still, lurkers' health literacy benefits from their exposure to the discussions. A model reflecting these insights (and extending the high-level research model of the present study, see Figure 2) is given in Figure 7.

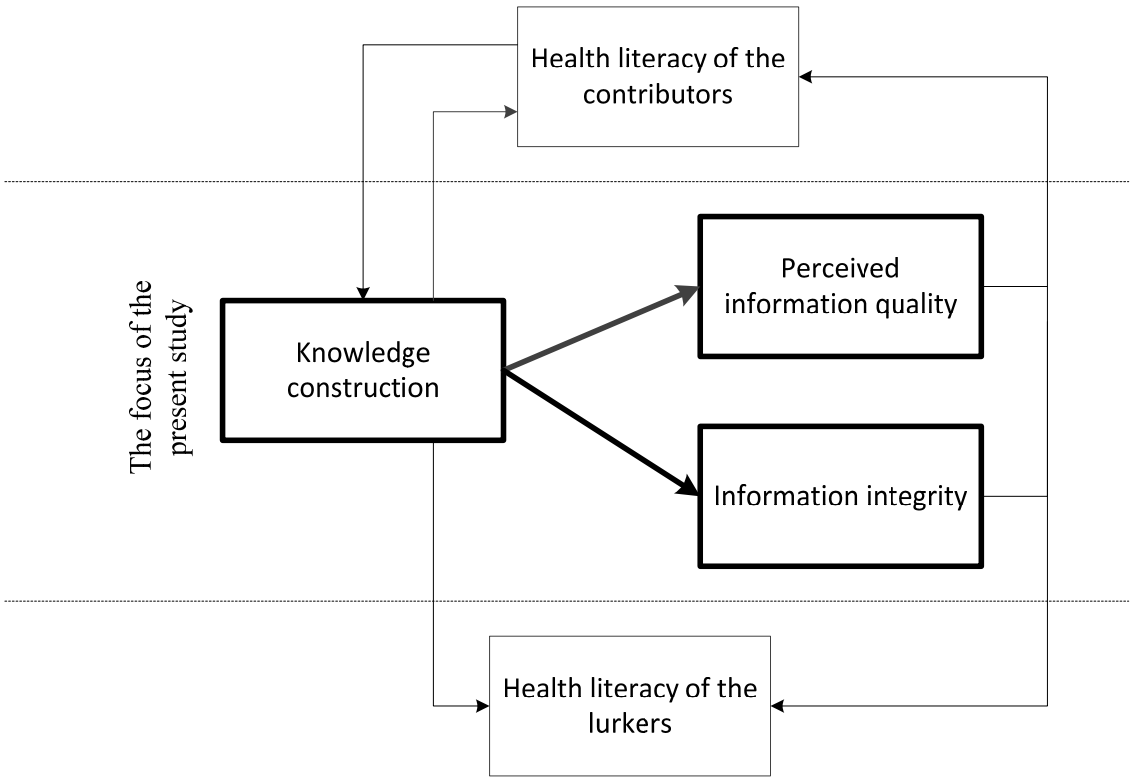

Figure 7. The focus of the present study and health literacy. The constructs and the hypothesis falling within the scope of the present study are shown with bold lines.

As highlighted in the diagram in Figure 7, the present study did not include health literacy. This was, in part, because health literacy is difficult to assess via content analysis, and the inclusion of health literacy would require a multi-method approach that was not feasible within the time and budget constraints of the present study. At the same time, the connections between health literacy, a matter of major practical concern,

and knowledge construction in health support group online discussions highlight the importance of pursuing the purpose of the present study.

2.12 Summary

This chapter reviewed the literature and formulated the knowledge gap addressed by the present study.

The ongoing transition to the patient-centred healthcare paradigm suggests that patients adopt a more active role in managing their conditions. Patient-centred healthcare suggests that patients become more knowledgeable, prompting patients to learn more about their health conditions. As the result, the Internet is becoming an important source of health-related information for patients. Internet-based health support groups are a resource allowing patients to both gain access to diverse information and diverse points of view and obtain moral support from people with similar problems. The chapter discussed in detail a review of experimental studies of the benefits of health-related peer-to-peer support groups; the results were inconclusive, which was attributed to small numbers of participants and to low participation in the discussions by the participants in the experimental groups (groups of participants that were expected to engage in online discussions). A study of motivation behind the participation in health-related online discussions was also discussed. The study concluded that accessing information was a stronger driver of discussion participation than seeking moral support.

Lurkers are individuals who read discussion forum messages, but never post any. Lurkers were characterized as important stakeholders (legitimate peripheral participants) who share knowledge constructed in the discussions in their off-line communities or apply the knowledge, thus enhancing the positive effect of the discussions.

Social constructivism suggests that knowledge is constructed in interactions between individuals engaged in substantive discussions. Social constructivism was the main theoretical foundation of this study.

To find a suitable (from the perspective of addressing the research questions of this study) approach to conceptualization and measurement of knowledge construction, the chapter presented a detailed review of existing studies involving coding transcripts of online discussions to detect and to measure knowledge construction. The studies were primarily descriptive: they described and characterized how knowledge construction occurs, rather than tested hypotheses regarding the effects or the antecedents of knowledge construction. I summarized and related the categories and levels of knowledge construction across studies, and demonstrated that most of the studies follow a similar pattern in categorizing knowledge construction activities, consistent with a view that there are lower and higher levels of knowledge construction. Moreover, I summarized the approaches to content analysis used in the studies and highlighted a number of discrepancies with the methodology suggested by Krippendorff (2004).

Knowledge construction at organizations can be viewed as part of knowledge creation—a central aspect of knowledge management. Knowledge construction at organizations can be viewed from the perspective of the scientific paradigm—as making sense of scientific knowledge in local contexts—or from the perspective of the social paradigm—as constructing knowledge locally, rather than importing knowledge.

The chapter presented in detail the findings of a descriptive empirical study of online discussions at a community of practice. The study was particularly of interest because it was conducted in a non-educational context. The study found that most of the members of the community of practice were lurkers; it also found that lurkers were active in terms of applying the knowledge they gained from the online discussions.

Information quality can be seen from the perspective of information integrity—the extent to which the information is objectively correct. An alternative perspective is perceived information quality—the extent to which the information is perceived as relevant, understandable, believable, useful, and so on. The chapter reviewed the conceptualization of perceived information quality by McKinney et al. (2002), which involved the dimensions of relevance, understandability, reliability, adequacy, scope, and usefulness.

In context of health support group online discussions, information integrity can be seen as the extent to which the recommendations for health-related behaviour or interventions represent evidence-based knowledge—scientific knowledge obtained via research designs minimizing bias. The chapter described the Cochrane systematic review—a process applied to assess research studies as presenting evidence-based knowledge.

The chapter concluded by summarizing the knowledge gaps discovered as the lack of studies of information quality of transcripts of online discussions, as well as the lack of studies of the effects of knowledge construction in online discussion (particularly, in non-educational contexts such as health support group online discussions).

Further, the chapter reviewed the literature on health literacy and highlighted the connections between health literacy and knowledge construction in health support group online discussions, presenting an argument that health literacy can be seen as both an antecedent and a consequence of knowledge construction. These connections further highlight the importance of pursuing the purpose of the present study.

Chapter 3. Model Development and Hypotheses

3.1 Introduction

Chapter 3 introduces in detail the conceptualizations of knowledge construction (as explicitation and evaluation), perceived information quality (as relevance, understandability, and usefulness), and information integrity. Then, the research model with knowledge construction and perceived information quality represented as their dimensions is introduced, along with detailed justifications of the research hypotheses.

3.2 Conceptualization of Constructs

Knowledge construction in online discussions happens via interactions between the discussion contributors. Even though each message contributes to the overall interaction, ultimately knowledge construction is not achieved in a single message, but over the whole interaction involving a sequence of messages. To test the high-level hypotheses suggested in section 1.6, there was a need to associate the quality of information presented by the discussion transcript with the level of knowledge construction. In the belief that the concepts of the level of knowledge construction and of the quality of information incorporate not only the content of individual messages, but also the message flow, I conceptualized these constructs as representing full discussions. To clarify this belief by using the metaphor of a Socratic dialogue introduced in section 2.8, my view is similar to a view that the wisdom and value of a Socratic dialogue can only be assessed by perceiving the dialogue in its entirety, rather than by mechanically adding up the wisdom and the quality of individual utterances.

While most of the studies introduced in section 2.6 associated aspects of knowledge construction with messages (13 of the 16 studies), there were three studies (by Newman et al., 1997, Järvelä and Häkkinen, 2002, and Puntambekar, 2006) that associated aspects of knowledge construction with the whole discussion thread, providing justifications similar to the one given above.

3.2.1 Knowledge Construction

The level of knowledge construction is the level of sophistication of a discussion, viewed from the point of view of the constructivist learning paradigm.

In conceptualizing the level of knowledge construction, this study follows the work by Veerman and Veldhuis-Diermanse (2001) and Schellens and Valcke (2005) and focuses on the distinction between explicitation, evaluation, and new information. I take a view that new information (such as a new fact, a new opinion, or a new theory) is always present in a thread that is task-related (devoted to a substantive issue, rather than a mere social exchange). Indeed, the presence of new information distinguishes such threads from the threads that are not task-related. Although new information is necessary to begin (and sometimes, to maintain) a discussion, the social constructivist learning starts to unfold only when the contributors go beyond simply posting new information to interacting with each other to discuss it, resulting in a pattern of messages involving explicitation and evaluation. Meaningful interactions may unfold in different threads to different degrees. Therefore, at the level of a discussion thread it does not make sense to consider the mere presence of evaluation and explicitation. Rather, it makes sense to consider the prevalence of explicitation and evaluation. Explicitation and evaluation constitute the collective "thought process" in the community of individuals networked together by the discussion, while new information is the material triggering the process.

Therefore, at the level of a thread, I conceptualize knowledge construction as a two-dimensional construct comprising two dimensions: explicitation and evaluation. The choice of explicitation and evaluation can be further supported by considering the outcome of the synthesis of coding schemes presented in section 2.6.6 (see Table 2-18). As argued in section 2.6.6.1, the categories of explicitation, integration, evaluation, and resolution can be seen as levels of knowledge construction. However, a model involving all of them would be too complex, and problems with discriminant validity would be likely because the distinctions between the neighbouring categories are not clear-cut. Explicitation and evaluation are spaced apart, having the integration category between them. At the same time, explicitation and evaluation are in the low to middle range of the scale—they cover knowledge construction activities that are not too sophisticated. This is appropriate for studying discussions in a non-educational setting,

where knowledge construction is not actively driven by an educator (see the discussion in section 2.6.2.1).

3.2.1.1 Explicitation

Explicitation is elaboration and refinement of earlier statements in the discussion. The definition of explicitation by Schellens and Valcke (2005) is given in Table 3-1. To further clarify the construct of explicitation, I extracted the keywords from the definition (*refining* and *elaboration*) and from the definition of evaluation by the same authors (from the part stating what evaluation is not—see Table 3-2) (*confirmation* and *negation*). A further keyword, *clarification*, was added from the definitions of explicitation by Henri (1992), Pena-Shaff and Nichols (2004), and Gunawardena et al. (1997). This resulted in the list of keywords in Table 3-1. The keyword definitions were taken from WordNet (a lexical database of English maintained by Princeton University in the US, http://wordnetweb.princeton.edu/perl/webwn).

Table 3-1 *The Content of the Explicitation Construct*

Definition by Schellens and Valcke (2005, p. 960)	Keywords
Type of communication that reflects further refining and/or elaboration of earlier ideas.	*refinement*—improving or perfecting by pruning or polishing. *elaboration*—developing in intricate and painstaking detail. *clarification*—an interpretation that removes obstacles to understanding. *confirmation*—additional proof that something that was believed (some fact or hypothesis or theory) is correct. *negation*—a negative statement or a statement that is a refusal of some other statement.

3.2.1.2 Evaluation

Evaluation is critical discussion of earlier information or ideas. The definition of evaluation by Schellens and Valcke (2005) is given in Table 3-2. To further clarify the construct of evaluation, I extracted the keywords from the definition (*critical discussion*, *argumentation, reasoning*, and *justification*). The keyword definitions were taken from WordNet.

Table 3-2 *The Content of the Evaluation Construct*

Definition by Schellens and Valcke (2005, p. 961)	Keywords
This type of written messages corresponds to a critical discussion of earlier information or ideas. It goes beyond a simple confirmation or negation and reflects argumentations, reasonings, [*sic*] justifications.	*critical discussion*—discussion characterized by careful evaluation and judgment. *argumentation*—a discussion in which reasons are advanced for and against some proposition or proposal. *reasoning*—presentation of reasons and arguments. *justification*—the act of defending or explaining or making excuses for by reasoning.

3.2.2 Perceived Information Quality

In conceptualizing perceived information quality, this study follows the work by McKinney et al. (2002). McKinney et al. conceptualized perceived information quality (which they called just "information quality") as six dimensions: relevance, understandability, reliability, adequacy, scope, and usefulness, with the items tested in three pilot studies (see section 2.8). For the reasons discussed further in this section, the present study did not include reliability, scope, or adequacy.

Reliability was defined by McKinney et al. (2002) as "concerned with the degree of accuracy, dependability, and consistency of the information" and was operationalized via "information is trustworthy", "information is accurate", and "information is credible" (p. 301). Reliability in the study by McKinney et al. is similar in content to accuracy by Wang and Strong (1996); Wang described accuracy as "data are certified, error-free, accurate, correct, flawless, reliable, errors can be easily identified, the integrity of the data, precise" (p. 14). In the present study, I took the view that reliability (particularly, when applied in context of a content analysis study) is close in content to information integrity, the degree to which a particular discussion thread promotes evidence-based rather than anecdotal knowledge (for an introduction of information integrity, see section 3.2.3). Even though the constructs are not the same, with information integrity referring to objective accuracy, and accuracy as part of the reliability dimension referring to perceived accuracy, I took the view that a separate study, focused on this particular distinction, is required to study it. Because information

integrity was to be included as a separate construct, I decided not to include reliability (or accuracy) in the model.

According to McKinney et al. (2002), scope "evaluates the extent of information, range of information and level of detail provided by the Web site" (p. 301). Nonetheless, the items they used to measure scope focused on variety and range, and not on the level on detail: "information covers a wide range", "information contains a wide variety of topics", and "information contains a number of different subjects" (p. 309). Both the definition and the operationalization of scope suggested that discussions that deviate from the focus on managing weight may have broader scope than discussions focused on weight management. At the same time, "level of detail" aspect was covered by the elaboration aspect of explicitation dimension of knowledge construction ("*elaboration*—developing in intricate and painstaking detail"—see Table 3-1), and breadth, variety, and range of arguments focused on managing weight were captured by confirmation and negation ("*confirmation*—additional proof that something that was believed, some fact or hypothesis or theory, is correct" and "*negation*—a negative statement or a statement that is a refusal of some other statement"). Therefore, I concluded that scope, as it was conceptualized and operationalized by McKinney et al., was not suitable for inclusion in the model because it does not take into account if the discussion is focused on weight management. When broader scope is achieved by conducting an unfocused discussion, it is likely to be lowering perceived information quality. However, when broader scope is achieved by conducting a rich discussion focused on weight management, the content of the scope construct appears to overlap with the content of the explicitation dimension of knowledge construction. Thus, the construct of information scope in part overlaps in content with other construct of the model and in part captures an aspect (how broad is the scope of off-topic discussions) that is not relevant to the purpose of the present study. Therefore, information scope was not included in the model.

Along with perceived information quality, McKinney et al. (2002) considered the dimensions of system quality: access, usability, entertainment, hyperlinks, navigation, and interactivity. The access dimension, described by McKinney et al. as "the speed of access and the availability of the Web site at all times" (p. 301) is similar to the accessibility dimension by Wang and Strong (1996): "accessible, retrievable, speed of

access, available, up-to-date" (p. 14). Even though Wang and Strong, unlike McKinney et al., present accessibility as a dimension of data quality, they do indicate that it is related to the system used to serve the data. In the present study, there was no difference between discussion threads in terms of speed of access or availability at all times, because all threads were served by the same system. Moreover, it is difficult to see how knowledge construction activities would influence accessibility—speed of access or availability are not affected by the content of messages posted by the discussion contributors. Therefore, accessibility was not included in the model.

Finally, adequacy refers to whether the information is enough, and just enough, to make a specific decision. Thus, adequacy is appropriate for assessing information provided in support of specific decisions, but not appropriate for assessing information generated in broader discussions.

Therefore, in this study the perceived information quality construct was conceptualized as a multidimensional construct comprising three dimensions (lower-level constructs): relevance, understandability, and usefulness. The dimensions of perceived information quality are listed in Table 3-3. The definitions of these constructs were initially given in section 2.8 and are repeated in the table. Keywords describing the content of each construct (dimension) were taken from the instruments McKinney et al. (2002) used to measure these constructs.

Table 3-3 *Dimensions of Perceived Information Quality*

Construct	Definition	Keywords
Relevance	extent to which information is applicable and helpful for the task at hand (based on Knight & Burn, 2005)	*applicable* *related* *pertinent*
Understandability	extent to which information is clear, without ambiguity and easily comprehended (based on Knight & Burn, 2005)	*clear in meaning* *easy to comprehend* *easy to read*
Usefulness	users' assessment of the likelihood that the information will enhance their ability to address their concerns (based on McKinney et al., 2002)	*informative* *valuable*

3.2.3 Information Integrity

Information integrity represents the degree to which a particular discussion thread promotes evidence-based rather than anecdotal knowledge. Thus, a discussion thread with high information integrity would promote recommendations that are rooted in medical research conducted without bias, with the lack of bias demonstrated in quality assessment processes such as Cochrane reviews. The concept of information integrity in application to medical information was discussed in detail in section 2.9.

3.3 Structural Model and Detailed Hypotheses

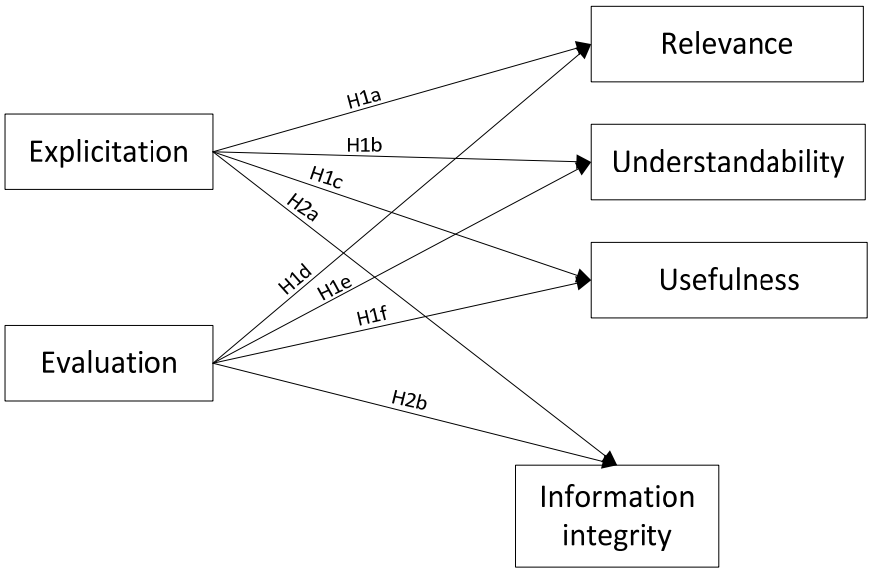

Figure 8. Research model from section 1.6 (Figure 2) with knowledge construction represented by explicitation and evaluation, and perceived information quality—by relevance, understandability, and usefulness.

The initial research model introduced in section 1.6 (Figure 2) was formulated for the constructs of knowledge construction, perceived information quality, and information integrity. The examination of these constructs in Chapter 2 and in section 3.2 suggested that knowledge construction and perceived information quality should be treated as multidimensional constructs. In this section the hypotheses suggested in Figure 2 are reformulated at the level of the individual dimensions of knowledge construction and perceived information quality (PIQ). The resulting hypotheses are consistent with the hypotheses suggested in section 1.6.

The overall model suggested by the hypotheses justified in this section is presented in Figure 8. In Figure 8, the low level hypotheses corresponding to the hypothesis H1 in Figure 2 are labelled as H1a, H1b, H1c, H1d, H1e, and H1f. The low level hypotheses corresponding to the hypothesis H2 in Figure 2 are labelled as H2a and H2b.

3.3.1 Knowledge Construction Results in Better Perceived Information Quality (H1)

This section introduces the hypotheses elaborating the hypothesis H1 introduced in section 1.6.

3.3.1.1 H_{1a}: Explicitation →PIQ—Relevance

Refinement, elaboration, and clarification make the new information meaningful in the context of the discussion by exposing the relationships between the new information and the pertinent aspects of the context. Information meaningful in a context is more likely to be perceived as applicable in the context or pertinent to the concerns of the contributors in the context. In health support group online discussions, elaboration and clarification of a new suggestion regarding health-related behaviour enable the contributors to see its relevance to addressing their health-related issues.

Based on the justification above, higher levels of explicitation create more opportunities for the information to be seen as relevant, so that ultimately explicitation causes relevance.

3.3.1.2 H_{1b}: Explicitation →PIQ—Understandability

Refinement, elaboration, and clarification result in the information being restated in easier to understand form that is more likely to be perceived as easy to read and easy to comprehend. In health support group online discussions, elaboration and clarification of new suggestions regarding health-related behaviour (e.g., restating the suggestions in easier to understand, day-to-day language in response to queries by participants not used to medical jargon) make it easier for the contributors to understand the information generated by the discussion.

Based on the justification above, higher levels of explicitation result in the information being reworded to make it easier to understand, so that ultimately explicitation increases understandability.

3.3.1.3 H_{1c}: Explicitation → PIQ—Usefulness

Refinement, elaboration, and clarification result in the new information being presented from a greater variety of perspectives, making it more likely that a perspective directly addressing the concerns of a particular contributor (and thus, valuable to the contributor) emerges in the discussion. Thus, explicitation makes it more likely that the information is perceived as useful. In health support group online discussions, elaboration of new information relating to a recommendation about a health-related behaviour is likely to result in the recommendation being restated in a form seen as informative and valuable by a particular contributor.

Based on the justification above, higher levels of explicitation create more opportunities for the information to be seen as useful, so that ultimately explicitation increases usefulness.

3.3.1.4 H_{1d}: Evaluation → PIQ—Relevance

Justification and critical discussion of propositions derived from the new information enable the contributors to see the relevance of the propositions. In health support group online discussions, propositions are likely to be in the form of recommendations for health-related behaviour or medical interventions.

Higher levels of evaluation create more opportunities for the propositions, such as generalizations, emerging in the discussion to be seen as relevant, so that ultimately evaluation increases relevance.

3.3.1.5 H_{1e}: Evaluation → PIQ—Understandability

Justification and critical discussion of propositions derived from the new information make it easier for the contributors to comprehend the propositions, as reasons behind the propositions are restated in easier to understand form to convince the contributors participating in the discussion. In health support group online discussions, critical

discussion is likely to prompt the restatement of propositions initially formulated using specialist medical terms in day-to-day language, thus making the resulting information easy to read and clear in meaning.

Based on the justification above, higher levels of evaluation result in information that is easy to understand being generated in the discussion, so that ultimately evaluation increases understandability.

3.3.1.6 H_{1f}: Evaluation → PIQ—Usefulness

Critical discussion of propositions derived from the new information results in the propositions being restated, reshaped, adjusted, and changed. The proponents of the propositions are likely to restate them in forms in which their usefulness is evident to other contributors. In health support group online discussions, propositions are likely to be in the form of recommendations for health-related behaviour or medical interventions. Critical discussion of such recommendations is likely to result in the recommendations being developed and restated to make them appear useful from the perspectives of the discussion contributors.

Based on the justification above, higher levels of evaluation create more opportunities for useful propositions to emerge in the discussion, so that ultimately evaluation increases usefulness.

3.3.2 Knowledge Construction Results in Better Information Integrity (H2)

This section introduces the hypotheses elaborating the hypothesis H2 introduced in section 1.6.

3.3.2.1 H_{2a}: Explicitation → Information Integrity

Elaboration of the new information may result in information lacking integrity being immediately negated. New information lacking integrity is exposed, and information integrity is thus preserved. In health support group online discussions, a new suggestion regarding health-related behaviour may be immediately rejected by contributors with sufficient background to recognize that the information is, in fact, not grounded in research.

Based on the justification above, higher levels of explicitation result in better assurance that information that is not evidence based is exposed as such, so that ultimately explicitation increases information integrity.

3.3.2.2 H_{2b}: Evaluation → Information Integrity

Critical discussion of propositions derived from the new information results in any flaws in the propositions being more readily discovered, leading also to the discovery of any problems with the integrity of the new information they are based on. In health support group online discussions, critical discussion is likely to result in recommendations that are counter to scientific evidence being rejected, thus resulting in higher integrity of the information generated by the discussion.

Based on the justification above, higher levels of evaluation result in better assurance that information that is not evidence based is exposed as such, so that ultimately evaluation increases information integrity.

3.4 Summary

This chapter introduced the conceptualizations of knowledge construction, perceived information quality, and information integrity. Knowledge construction was conceptualized as a multidimensional construct with the dimensions of explicitation and evaluation. Explicitation describes the extent to which lower level knowledge construction activities are present in the discussion, and evaluation describes the extent to which higher level knowledge construction activities are present. Perceived information quality was conceptualized as a multidimensional construct with the dimensions of relevance, understandability, and usefulness. Information integrity was conceptualized as the extent to which recommendations regarding health-related behaviour or interventions in the discussion reflect evidence-based knowledge.

In the research model both explicitation and evaluation were hypothesized to affect information integrity and all of the dimensions of perceived quality.

Chapter 4. Research Methodology

4.1 Introduction

This chapter describes the research methods used in the present study.

First, the overall approach to research is stated as positivist, quantitative content analysis. Then, the approach to content analysis employed in this research is characterized from the perspectives of the criteria and attributes suggested by Krippendorff (2004), including the Krippendorff's criteria for suitability of content analysis as a technique, the Krippendorff's framework for content analysis, and the Krippendorff's criteria for data languages.

Next, the choice of the content source (a health support group devoted to weight management) and this study's approaches to unitizing, sampling, screening, and coding are discussed. The operationalisations of the constructs are also introduced and justified.

The chapter continues by discussing the steps taken to assess and ensure the reliability of coding and the reliability of measures of latent constructs, as well as the steps to assess and to ensure validity. Validity is considered from the perspectives of face, content, nomological, construct, convergent, discriminant, internal, external, ecological, and social validity.

The chapter concludes by introducing the statistical techniques used for testing the research model (the measurement model, in terms of convergent and discriminant validity, and the structural model). The properties of data required to ensure the validity of the statistical tests employed in this study are also discussed.

In this chapter I follow the recommendations of the APA Manual (American Psychological Association, 2010, p. 116), and do not provide references for statistical procedures or for methodology-related concepts and practices in common use.

4.2 Overview of Research Procedures

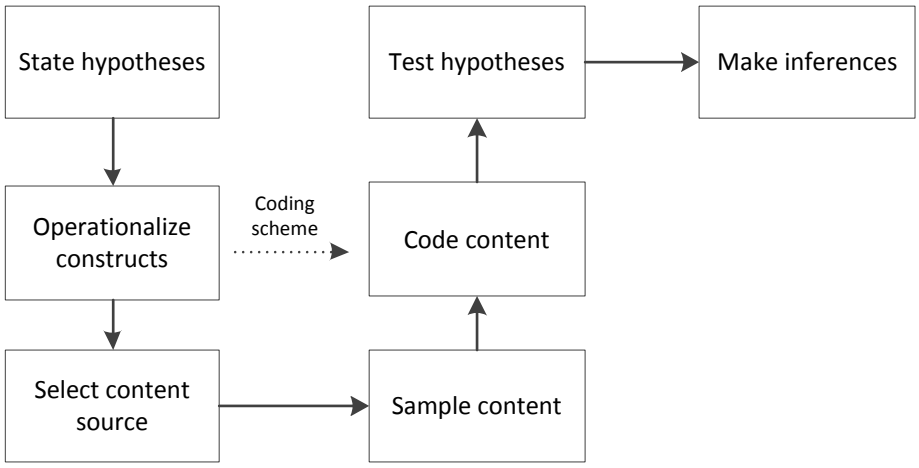

Figure 9. The main steps of this study.

Figure 9 provides an overview of the main steps involved in this study. The details of the overall approach and of each step are discussed in the rest of this chapter.

4.3 Overall Approach to Research

This section discusses the overall approach to research adopted in this study.

4.3.1 Positivism Versus Interpretivism

Positivist research involves formulating propositions (hypotheses) based on theory. Positivism suggests that truth is objective and that the truth of a proposition can be tested against data (Orlikowski & Robey, 1991). In tests against data, a proposition can be proved to be wrong, but a proposition cannot be proved to be correct. A proposition that was not demonstrated to be false in more and in a greater variety of tests is seen as supported by data and is treated as being true. The results of positivist research can be generalizable: the researcher may be able to state the circumstances in which the propositions tested in the research are expected to be true, independent of anyone's subjective judgment.

Interpretivist research aims at achieving an understanding of the social world from the perspectives of the researchers and of the research participants. Interpretivist research emphasizes social interaction, language, consciousness, and shared meaning (Boland, 1985; Myers, 1997). Interpretivism does not assume the existence of an objective truth, but rather allows multiple perspectives with respect to the same social reality to coexist; rather than trying to determine the "true" perspective, an interpretivist researcher attempts to capture and to faithfully report the perspectives of all important stakeholders. Rather than claiming to avoid bias due to subjective judgments, an interpretivist researcher recognizes the possibility of bias in interpreting the results and provides sufficient information for the reader of the research to understand the nature of the possible bias. Interpretivist research does not claim generalizability; it is up to the reader of the research to decide if the research outcomes are applicable to the reader's context. Some of the methodologists describe interpretivist research in terms of knowledge construction—knowledge is constructed in a dialogue between the researchers and the research participants (Kvale, 1995). Yet, the emphasis on knowledge construction in interpretivist research is not universal; for example, the highly influential introduction to grounded theory by Corbin and Strauss (2008) relies very little on the knowledge construction metaphor (although the knowledge construction view of interpretivist research is positively mentioned in the introduction, Corbin and Strauss, p. 10).

In this study the research questions were formulated as hypotheses, thus suggesting the use of the positivist paradigm. The hypothesized effects of knowledge construction on perceived information quality and on information integrity were tested against empirical data.

The positivist research stance adopted in the research method relates to formulating the research design of the present study. At the same time, I adopt a social constructivist view of discussion forum dynamics (see section 2.5) and I rely on knowledge construction as a concept capturing the important aspects of such dynamics (see sections 1.5, 2.7, and 3.2.1). Thus, I take a view that describing discussion forum dynamics in terms of knowledge construction, with the resulting transcript reflecting the knowledge construction dynamics, is a useful and widely accepted way to make sense of such dynamics. However, I do not take an interpretivist approach to studying

knowledge construction, but rather operationalize the knowledge construction construct to measure it and to test hypotheses involving it.

There is no contradiction in relying on the knowledge construction metaphor in describing the target phenomenon, and not relying on the same metaphor in formulating the research design—modelling the target phenomenon and modelling the research design are two separate issues. Indeed, the richest understanding of a given social phenomenon is likely to be achieved by addressing the phenomenon by a variety of methods, relying on different paradigms (Creswell, 1998). One can rely on interpretivist research to understand the rich perspectives of the contributors on knowledge construction in health support group online discussions. This would offer a rich and deep understanding of the phenomenon, but the research would be vulnerable to claims that the results lack objectivity. Positivist research, including the approach adopted by the present study, is reductionist—complex interactions in the discussion are projected onto a small set of variables, and the richness of the data is not fully taken advantage of. The results, however, can be seen as more objective and thus having greater relevance beyond the immediate research participants. In the long run, it is desirable that knowledge construction in health support group online discussions is studied by a variety of methods; this, however, is hardly feasible to achieve in a single study. In the present study, I opted for a positivist approach in an expectation that if the results are found interesting and significant by the broader research community, interpretivist studies would follow.

A very similar stance was adopted by Schellens and Valcke in two studies. Schellens and Valcke (2005), in a study of educational science freshmen contributions to a discussion group, confirmed the hypothesis that higher levels of participation result in higher levels of knowledge construction. Schellens and Valcke (2006), in another study of educational science freshmen contributions to a discussion group, confirmed the hypothesis that at the end of the discussion period the levels of knowledge constructions are higher than at the beginning (and thus, the contributors improve their knowledge construction skills over the discussion period). These studies, similarly to the present study, conceptualized technology mediated interactions in terms of knowledge construction, operationalized knowledge construction, and tested hypotheses regarding relationships of knowledge construction to other constructs. Some

further, recent studies that have conceptualized discussion forum dynamics as knowledge construction, measured knowledge construction, and statistically tested hypotheses regarding how knowledge construction relates to other constructs (and thus, relied on positivism in research design, but conceptualized discussion forum dynamics in terms of social constructivism) are the studies by De Wever, Van Keer, Schellens, and Valke (2009, 2010) and by Rienties et al. (2012) discussed in section 6.3.1.1.

4.3.2 Quantitative Research Versus Qualitative Research

Quantitative research involves representing social reality using quantities and mathematical relationships. Quantitative research is often associated with the positivist paradigm because quantitative research offers a way to test hypotheses—by applying inferential statistics to measured quantities. In quantitative research hypotheses regarding causal relationships can be investigated by applying inferential statistics to test relationships between variables (Sale, Lohfeld, & Brazil, 2002).

Qualitative research involves working with rich data, such as texts and observations, which allows researchers to construct new understandings of the underlying social reality. Qualitative research is primarily used in conjunction with the interpretivist research paradigm (Sale et al., 2002).

This study relied on quantitative research to test the research hypotheses stated in section 1.6. In this study, the effects of knowledge construction on perceived information quality and information integrity were investigated by operationalizing the constructs as variables, followed by using inferential statistics to investigate the relationships between these variables.

4.3.3 Explanatory Research Versus Exploratory Research

Exploratory research involves investigating little-understood phenomena and generating hypotheses for further research (Marshall & Rossman, 2010). Exploratory research is conducted in areas where the theory is either not available or is weak.

Explanatory research focuses on testing theories (Marshall & Rossman, 2010). Explanatory research is conducted in areas where there is sufficient understanding of social phenomena, and thus the theory is relatively strong.

This study was to a large degree exploratory, because it addressed the effects of knowledge construction on perceived information quality and information integrity; no study has attempted this in the past.

4.3.4 Approach to Data Collection

The most common approaches to data collection in quantitative research are experiment, survey, and content analysis.

An experiment requires changing the values of independent variables to compare outcomes (i.e. changing the level of knowledge construction to observe the effects of knowledge construction on information quality). Researchers modify the values of independent variables to observe the changes in response variables. In the present study, the use of an experiment as a research technique was not feasible because of the difficulty in controlling the level of knowledge construction. Because discussions at communities of practice occur naturally, and communities of practice are self-regulating, an attempt to control such a discussion would destroy the ecological validity of the study.

Both surveys and content analysis include studying populations, but in surveys the populations are of human participants, and in content analysis the populations are of documents. Information quality is a property of a document, and the level of knowledge construction, as conceptualized in this study, is a property of the dialogue captured in a document. Neither information quality nor knowledge construction can be assessed by surveying the document contributors. Therefore, a survey was not an option.

Content analysis involves sampling documents, coding documents to measure constructs (such as the dimensions of knowledge construction, perceived information quality, and information integrity), and testing hypotheses regarding relationships between the constructs. As content analysis involves working with authentic documents, the results of content analysis have high ecological validity. Moreover,

content analysis was feasible to execute within the time and financial constraints of this study.

In view of the considerations presented in this section, this study used content analysis.

One drawback associated with using content analysis is that, unlike experiments, content analysis does not distinguish causes from effects. However, as in this study information is generated via discussion (and not the other way round), it appears more plausible that various aspects of the discussion (such as the level of knowledge construction) affect the quality of the resulting information, rather than the other way round.

Compared to collecting data from the contributors directly (as in a survey), the drawback of using content analysis was that ultimately the results of the content analysis were used to make abductive inferences about the context of the content generation and use. These contexts involved both contributors and lurkers. Thus, more insightful results could have been obtained by combining content analysis with a survey, as in Newman et al. (1997) and Kanuka and Anderson (2007). However, this was not possible, as contact details of the initial contributors (and, particularly, of lurkers) were not available. Moreover, there were concerns that disrupting discussions by including contributors within the scope of the research would result in harm. Therefore, the option of combining content analysis with a survey was considered, but rejected, and the research was limited to conducting content analysis.

4.4 Approach to Content Analysis

This section introduces the approach to content analysis used in this study in context of the recommendations by methodological literature.

4.4.1 Reaffirming the Suitability of Content Analysis

Content analysis is a research technique for making valid inferences from texts to the contexts of their creation and use (Krippendorff, 2004). Online discussions are social interactions intermediated by text; the interaction aspect of online discussions is fully

captured in the resulting transcript (assuming the discussions are not accompanied by off-line interactions).

Content analysis is particularly suitable as a research technique when the social phenomenon that generated the content is repetitive, routine, and public, rather than rare and unconventional (Krippendorff, 2004). Discussion forum discussions follow repetitive patterns (suggested and, to an extent, enforced by the discussion software). The protocol involved is routine for repeat contributors, and the discussions are conducted in the public domain. In this study, the discussions were globally public, that is, they were open for access to anyone with access to the Internet, were not password protected, and were discoverable via search engines. Thus, according to the criteria suggested by Krippendorff, content analysis was an appropriate technique.

In this study, content analysis of publicly available transcripts of online discussions in a health support group was aimed at exploring relationships between constructs relevant to three different contexts of use: (a) the context of content contribution (knowledge construction), (b) the context of content consumption (perceived information quality), and the context of the effect on the consumers of the content (information integrity).

Krippendorff (2004, p. 82) suggests that before choosing content analysis as a research technique, the following questions need to be answered: why the available texts came into being, what they mean and to whom, how they mediate between antecedent and consequent conditions, and whether they enable the analysts to select valid answers to questions concerning their contexts. The rest of this section analyses the approach adopted in this study from the point of view of these questions.

Why the available texts came into being? Discussion contributors in health support group online discussions need to address health or appearance issues. Thus, they need both actionable advice and moral support. The contributors share interest in the same topic and use the discussion to mutually satisfy their needs. The contributors remain anonymous, and thus can frankly discuss issues. The contributors can contribute irrespective of time or place.

What they mean and to whom? From the point of view of the contributors (and lurkers), these texts are a source of (a) actionable information and (b) moral support.

From the point of view of medical doctors and the society at large, these texts could prompt their consumers to engage in both health enhancing and unhealthy behaviours, depending on the integrity of the advice they contain.

How they mediate between antecedent and consequent conditions? Contributors, as a community of practice, build knowledge relevant to weight loss (deposited in their minds, in connection between the contributors, as well as in the discussion transcript). Thus, the content mediates from a state of little knowledge in the system to a state with greater knowledge.

Do the texts enable to select valid answers to questions concerning the texts' contexts? The texts capture both the knowledge constructed in discussions and the knowledge construction process involved in knowledge construction. Therefore, both knowledge construction and the attributes of knowledge (such as its integrity) can be measured. Therefore, hypotheses regarding relations between these constructs can be tested.

4.4.2 Framework for Content Analysis

Krippendorff (2004) suggested a framework for content analysis involving a body of text, a research question, a context to make sense of based on the body of text, an analytical construct relevant to the context, inferences used to answer the research question, and validation of evidence. Table 4-1 describes the approach adopted in the present study in terms of the Krippendorff's framework.

As seen in the table, all of the components of the Krippendorff's framework were clearly present in the present study, suggesting that the approach used in this study was sound and complete.

4.4.3 Data Making

According to Krippendorff (2004, p. 83), data making involves reducing raw unedited texts to data that can be used for making inferences and includes unitizing, sampling, and coding. In this study, unitizing was addressed by choosing a discussion thread as a

unit of analysis (see section 4.6), the sampling procedure involved retaining only typical threads and random selection (see section 4.7), and the coding procedures were based on using independent coders for constructs hypothesized to be related, with two coders per construct to assess inter-rater reliability (see section 4.8).

Table 4-1 Approach to Content Analysis in the Present Study in Terms of Krippendorff's (2004) Framework for Content Analysis

The Krippendorff's framework	The present study
A body of text	The texts were transcripts of online discussions of an online health support group.
A research question that the analyst seeks to answer by examining the body of text	The research questions were stated as hypotheses involving the constructs of the level of knowledge construction, perceived information quality, and information integrity.
A context to make sense of based on the body of text	Knowledge construction is relevant to the context in which the transcript is generated (the actual discussion); perceived information quality and information integrity are relevant to the contexts of use of the transcripts (for example, by lurkers).
Inferences used to answer the research question	The relationships between knowledge construction, perceived information quality, and information integrity have immediate implications for how knowledge construction activities in the context in which the transcript is generated affect the contexts of use by promoting or failing to promote healthy behaviours.
Validating evidence, which is the ultimate justification of the content analysis	The dimensions of the target constructs were treated as latent variables, with reflective indicators assessed via a coding scheme. The hypotheses were tested by using linear regression and structural equation modelling against the body of text. The validity was considered from multiple perspectives, which are detailed in section 4.11.

4.4.4 Defining the Semantics of Data

To test the hypotheses of the study (introduced in section 3.3), the analytical constructs of explicitation and evaluation (dimensions of knowledge construction), and relevance, usefulness, and understandability (dimensions of perceived information quality) were treated as latent variables. Each latent construct was measured via a set of indicators

associated with it. Information integrity was coded directly (in the subsequent discussion, information integrity is treated as its own indicator).

Indicators were coded on semantic differential scales. As coders did not need to assign coding units to categories, exhaustiveness or mutual exclusivity were not issues to be addressed. In using semantic differential scales in context of content analysis, this study followed an approach demonstrated to be effective by Rourke and Anderson (2004).

Even-point (8 point) scales were used to encourage the research assistants acting as coders to reflect on each item before choosing the code (with the exception of the items used for the dimensions of perceived information quality, for which 11 point scales were used to maintain consistency with prior research). Coders were encouraged to contact the lead researcher if they found that a reasonable choice could not be made; thus, forcing the coders to make a choice (by using an even-point scale) when a choice could not be made was not an issue. Coders did not report problems for any of the items.

The details of how constructs were operationalized are discussed in sections 4.8.4 (for the dimensions of knowledge construction) , 4.8.5 (for the dimensions of perceived information quality), and 4.8.6 (for information integrity).

4.4.5 Operationalization as Data Language

Krippendorff (2004) considered operationalization from the perspective of defining a data language—a formal language representing the relevant aspects of the content under study.

In the present study, the data language was structured as a set of variables (indicators) measured by using semantic differential scales (to distinguish these variables, which were measured directly, from variables corresponding to latent analytical constructs, I refer to them as "manifest variables" or "indicators").

Krippendorff (2004, p. 151) suggested three criteria for data languages. The following three paragraphs restate the criteria and discuss how they were met in the present study.

The data language must be free of ambiguities and inconsistencies. In the present study, to ensure that there were no ambiguities, the coding scheme was reviewed both by other researchers and the coders, and coders were encouraged to provide feedback if inconsistencies or ambiguities were discovered during the coding process.

The data language must satisfy the requirements of the analytical techniques to be used. The present study relied on Structural Equation Modelling (SEM) for data analysis and hypothesis testing. The data language fulfilled the requirements of SEM to provide a set of indicators reflecting the values of each latent variable. Sufficient numbers of points were employed in the scales to allow treating the indicators as measured on an interval scale—a property desirable for SEM analysis.

The data language must transmit enough information about the phenomena of interest. For each construct, the set of items measuring the construct was designed to capture all of the relevant essential content of the construct (thus ensuring content validity). As much as possible, items were based on measures validated in prior research.

4.5 Choice of Content Source

A health support group online discussion site devoted to weight management was used as a source of content. The site was chosen because the discussions were substantive enough to suggest that knowledge construction was happening and because the transcripts were available in the public domain without restrictions (e.g., not password protected and discoverable via search engines).

The discussions that generated the content were asynchronous. Asynchronous discussions are likely promote higher levels of learning, because the contributors have sufficient time to reflect on their messages and can access the history of the discussion (Pena-Shaff & Nicholls, 2004).

4.6 Unitizing

Krippendorff (2004) distinguished sampling units, coding units, and context units. A sampling unit is sampled for inclusion in the analysis, a coding unit is a part of the sampling unit with which codes are associated, and a context unit is the body of text taken into account when associating codes with a coding unit.

In this study sampling, coding, and context units were delimited as follows.

The sampling unit was a discussion thread. Criteria for the inclusion of discussion threads for analysis are discussed in section 4.6.

For the dimensions of level of knowledge construction and perceived information quality the coding unit was also a discussion thread.

For information integrity, the coding unit was a recommendation for health-related behaviour or for an intervention. When extracting recommendations, coders took into account the content of the whole thread to make sure that the recommendation was indeed suggested by the overall discussion (and not tentatively stated, but later rejected). Thus, for the unitizing process the context unit was the whole thread. Nonetheless, the expert coders who coded the recommendation for information integrity did so solely based on the text of the recommendations. Thus, for coding recommendations, a recommendation was both the coding and the context unit. (This was because of resource limitations—the expert coders had sufficient time to read and to code recommendations, but not to read the full text of the threads.)

Operationalization of constructs and coding procedures is discussed in section 4.8.

Krippendorff (2004) suggested that greater consistency between coders can be achieved if the coding unit is small and simple in structure. Even though in many studies of knowledge construction the unit of analysis was a message, as discussed at the beginning of section 3.2, I believe that a message is not appropriate to judge the level of knowledge construction because at the level of a single message the overall interaction is not represented. Similar considerations apply to information quality. Thus, even though by using a smaller coding unit, such as a message or even a sentence, one can achieve greater reliability, this comes at the expense of validity. Therefore, I chose the

approach to unitization to assure the validity of the results, rather than to maximize reliability.

This study used a thread as a coding unit for knowledge construction because in social constructivism the learning process takes place in interaction between individuals (Leahey & Harris, 1985; Salomon, 1993). Furthermore, interaction, communication, co-construction, and negotiation are only reflected at the level of the discussion thread (Stahl, 2004; Weinberger & Fischer, 2006). Contributors interact with each other, and at the level of a thread one can see not just a single conversational turn, but the whole of the interaction process. Dennen (2008) emphasized the need to use large enough coding units in studying the process of knowledge construction by stating that knowledge is not series of discrete facts, and therefore documentation of learning through coding and counting small units in isolation is insufficient.

4.7 Sampling and Screening

The present study sampled from the threads initiated between the first week of August 2010 and the third week of October 2010 (1,390 discussion threads overall). At the time when the data were collected, all of these threads were inactive for over at least one month.

The population was narrowed by removing from consideration the threads falling outside a specific range for the following parameters: the number of messages and the number of contributors. This was done to focus on threads meeting the following conditions: (a) the threads should be large enough to allow knowledge construction and (b) the threads should be small enough to allow coders to deal with each thread as a single unit. In very small threads, there is little, if any interaction, and, therefore, the concept of knowledge construction is not applicable. On the other hand, in very large threads, the number of messages is too large for a coder to remember them at a time; moreover, in a very large thread the nature of discussion is likely to shift over time (for example, an on-topic thread may eventually turn into an off-topic thread, or even into a flame war). Thus, it may be difficult to code very large threads as single units, and the validity of the resulting codes may be in doubt.

Removing the threads that are too small or too large is similar to removing outliers. Outliers are observations that are so different from the rest that they are likely to be aberrations not representing the population (Hair, Black, Babin, & Anderson, 2009). By retaining only typical threads, this study went beyond the common practice of removing outliers, as usually only very small numbers of cases are removed to ensure that the remaining cases still represented the population (Hair et al., 2009). Because a relatively large number of threads were removed in the present study, the results did not necessarily generalize to all threads, but rather to the "typical" threads as defined in this study. I do not believe that this restriction undermines the ability of this study to address the stated research questions. Under any circumstances, it is not possible to sample from the population of all discussion threads in existence. It is sufficient to test the hypotheses for a class of threads that are representative (in an analytical rather than a statistical sense) of a typical thread in a typical discussion. I believe that in terms of achieving this, not including all threads in the discussion, but instead only threads that are not too large or too small, does not undermine the internal or the external validity of this study (validity is discussed in sections 4.11.5 and 4.11.6).

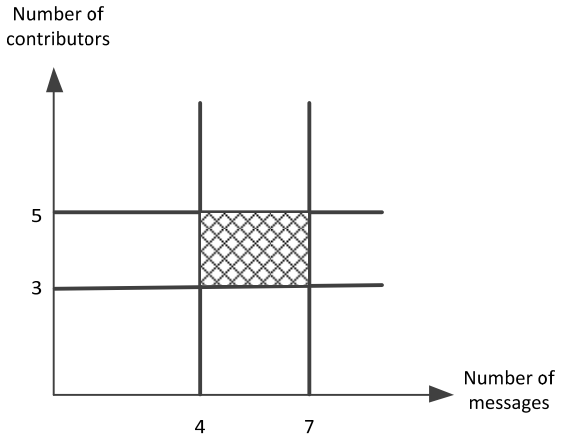

Figure 10. Definition of "typical" threads (given by the shaded area).

The parameters used to establish cut-offs and the exact cut-off values were based on my (subjective) observations and correspond to what I believed to be a "typical" thread size. Threads with the number of messages from four to seven and the number of contributors from three to five were retained. This roughly corresponded to threads falling within the second and the third quartile of the distribution of threads by the

number of messages and by the number of contributors (the shaded area in Figure 10). As the result, 501 threads were judged as typical and retained for further consideration. The following three paragraphs consider the rationale and the consequences of using these cut-off points in more detail.

Number of messages. Of the threads that were removed, 361 threads contained three or fewer messages (arguably, too short for an argument to develop—indeed, a thread consisting only of one or two messages can hardly be called a thread), and 418 threads contained eight or more messages. As seen in the histogram in Figure 11, there was a sharp drop between the numbers of threads of length seven and of length eight, suggesting seven as a cut-off value.

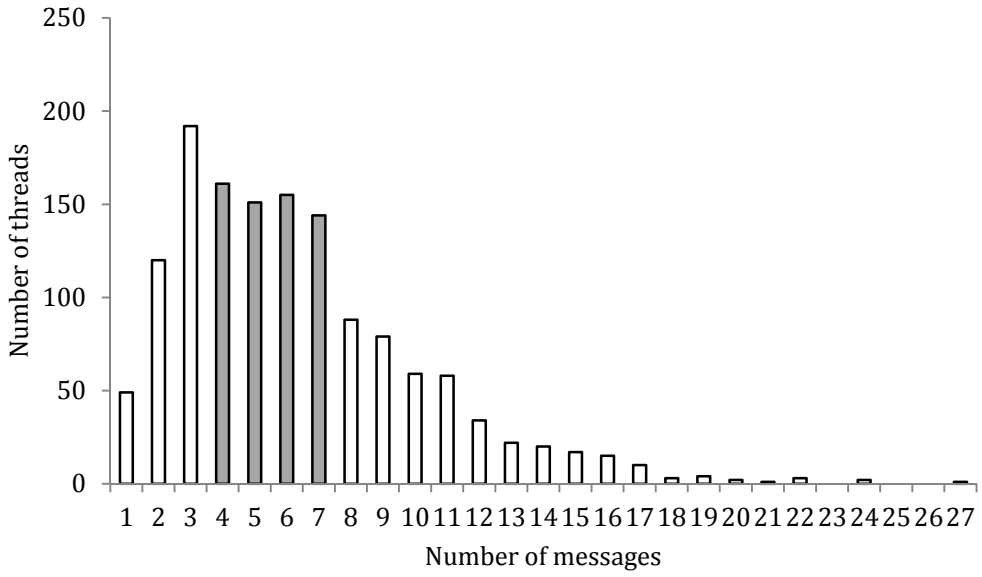

Figure 11. Distribution of threads by the number of messages in a thread. Bins corresponding to typical threads are shown in grey.

Number of contributors. Of the threads that were removed, 345 threads involved just one or two contributors (arguably, not enough to present multiple points of view), and 298 threads contained six or more contributors. As seen in the histogram in Figure 12, there was a sharp drop between the numbers of threads that involved five contributors and that involved six contributors, suggesting five as a cut-off value.

An alternative to setting cut-off values based on my subjective observations would be to conduct a separate content analysis study (involving coders) to distinguish threads that are too small for the argument to develop and long threads with the nature of the discussion shifting over time. This would necessitate the development of a coding scheme designed to distinguish such threads, and the coding scheme would have to be developed from scratch—to the best of my knowledge, a coding scheme of this type has never been developed. Moreover, applying such a coding scheme would considerably increase the costs of conducting the present study, because threads of all sizes would need to be coded. In terms of time and funding available for the present study, it was not feasible to use an extra round of content analysis to distinguish typical threads.

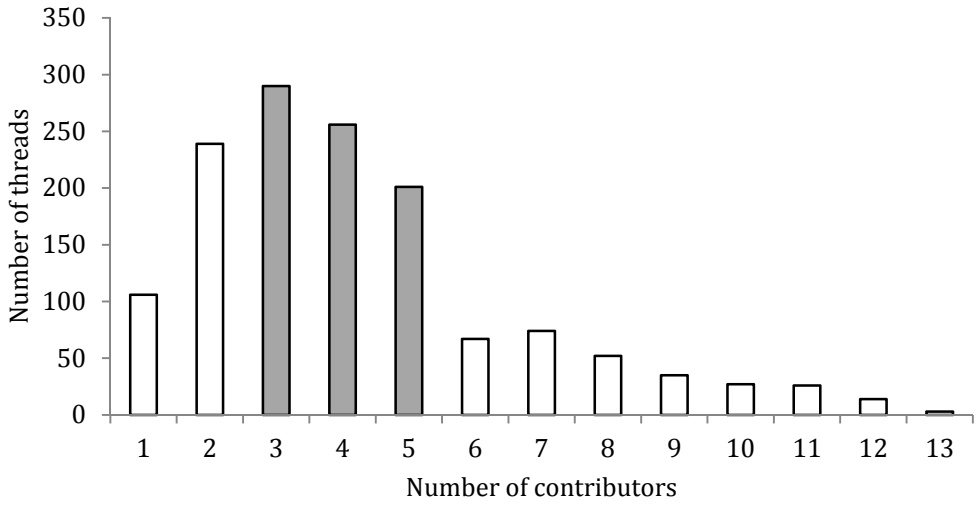

Figure 12. Distribution of threads by the number of contributors. Bins corresponding to typical threads are shown in grey.

After distinguishing typical threads, out of the pool of the 501 typical threads, 120 threads were drawn randomly for analysis, following the standard sampling procedure used in content analysis. Because the threads were chosen at random, they represented the population of typical threads—drawing threads at random did not introduce bias. The number of threads selected at random for analysis (120 threads out of 501 typical threads available in the sample) was based on two considerations. First, the budget of the study was limited, and hiring coders to code 120 threads was within budget. Second, the statistical techniques used to test the structural model in Figure 8 required approximately 100 cases (see section 4.13 for a more detailed discussion of this aspect).

The reduction in the number of threads reduced the statistical power of the study; however, because statistically significant relationships were detected, and various reliability and validity criteria were met (see section 5.6), this number was sufficient.

The screening of threads is summarized in Table 4-2 in terms of the number of threads and the number of messages retained after each phase.

Table 4-2 *Screening of Threads*

Screening phase	Number of threads	Number of messages
Threads initiated between August and October 2010	1,390	8,816
After extracting typical threads	501	2,732
After random selection	120	708
After data extraction	108	633

The 120 discussion threads obtained in random selection were inspected by two independent coders for the presence of specific recommendations addressing the issue of weight control. The details of the inspection procedure are given in section 4.8. Recommendations for weight control were discovered in 108 threads. These threads were treated as task related and used for further data analysis.

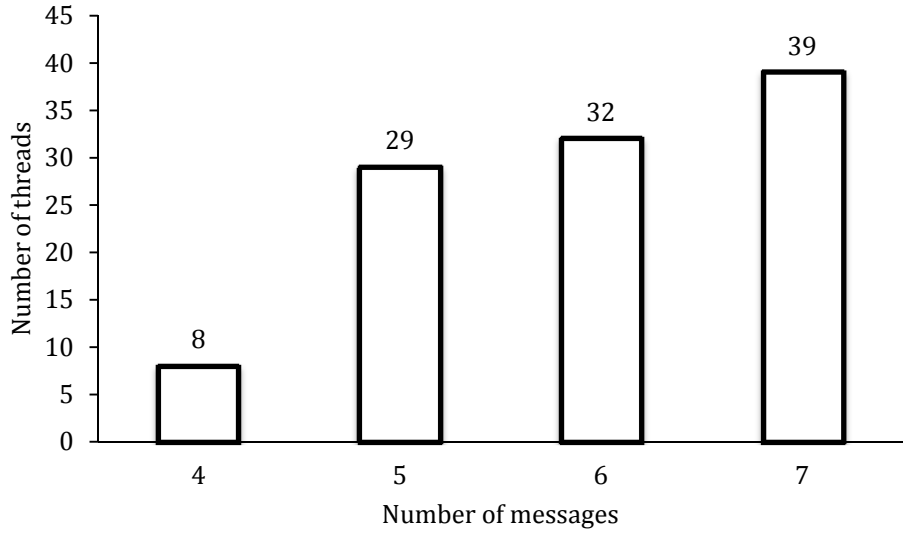

Figure 13. Distribution of task related threads by the number of messages in a thread.

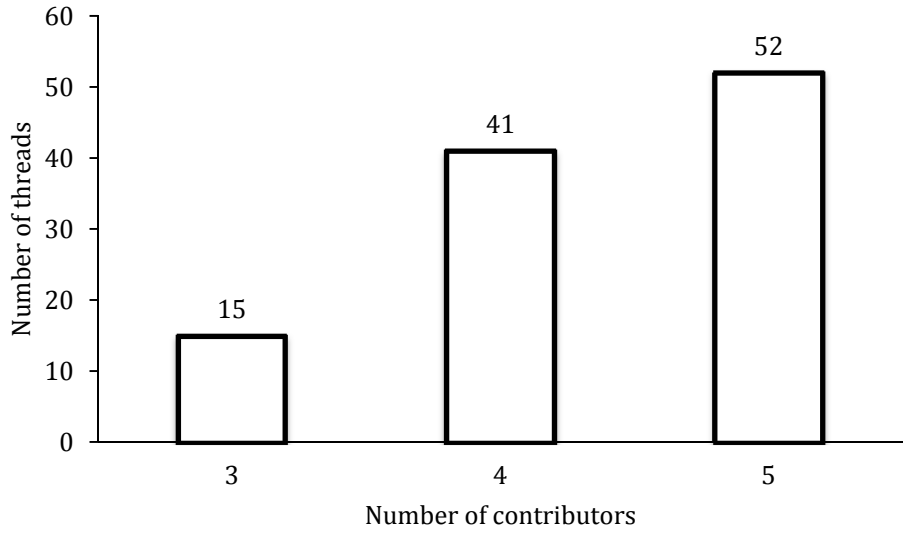

Figure 14. Distribution of task related threads by the number of contributors.

Histograms for the distributions of threads by the number of messages and by the number of contributors for the 108 task related threads that were coded are given in Figure 13 and Figure 14.

4.8 Coding

This study relied on coders (suitably qualified research assistants) to assign values to analytical constructs (dimensions of knowledge construction and perceived information quality, and information integrity) used to test the hypotheses (section 1.6) to achieve the purpose (section 1.2) of this study. The constructs could not be measured by using mechanical computation; therefore, it was necessary to involve coders and to rely on their judgments. As the study relied on the positivist research paradigm, it was assumed that the human judgments of the coders reflected an objective reality. That is, it was assumed that there were some objective, coder-independent values of the dimensions of knowledge construction and perceived information quality, and information integrity; coder judgments (imperfectly) reflected this objective reality.

Coding instructions (see Appendix A) were developed to (a) specify the recording process for the coders, to (b) assure that the meanings of the results can be understood by others, and (c) to enable the replication of this research by other researchers. To ensure that (b) and (c) are achieved, coder training was minimal, and the coders did not interact with each other and worked primarily based on the coding instructions presented in Appendix A, rather than based on undocumented training sessions or undocumented (and difficult to replicate) interactions between coders.

4.8.1 Overview of Coding Procedures

The coding procedures are outlined in Figure 15. I specified coder qualification (as discussed in section 4.8.2) and employed suitable coders. To minimize the risk of common method bias (discussed in section 4.8.2), large numbers of coders were employed, so that constructs hypothesized to have a cause effect relationship are never coded by the same coder, so that the outcomes of hypotheses testing reflect objective relationships rather than coder interpretations.

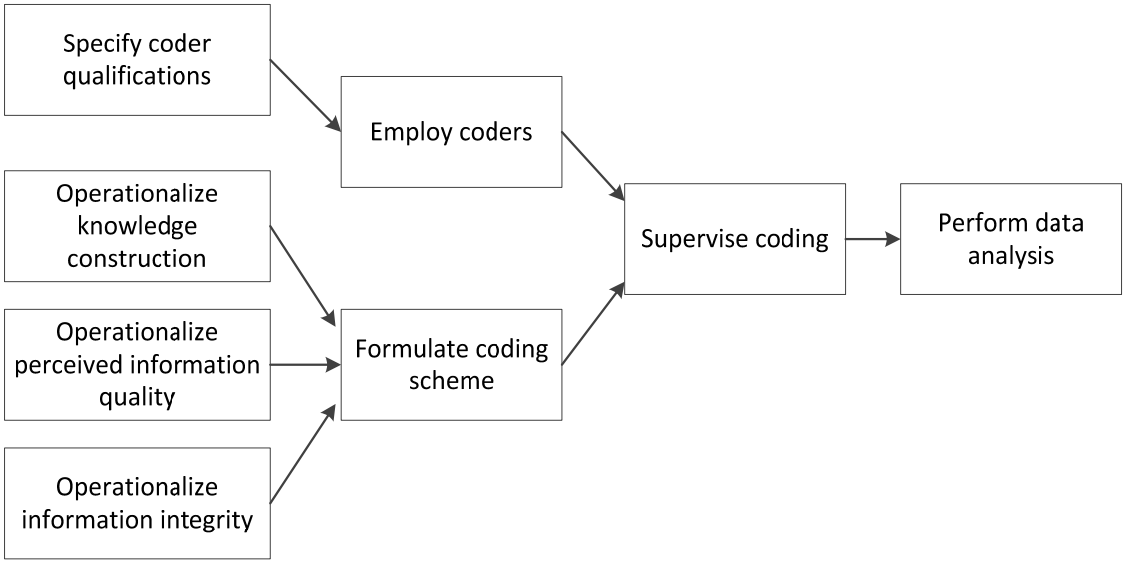

Figure 15. Coding procedures.

For perceived information quality, I was able to reuse an operationalization developed in prior research (see section 4.8.5). For knowledge construction, I developed an operationalization by synthesizing the approaches used in prior literature (see section 4.8.4). Finally, for information integrity, I developed an operationalization based on the Delphi method (see sections 4.8.6 and 4.8.7). Using these operationalisations, I developed a set of coding schemes (see Appendix A, Appendix B, Appendix C, and Appendix D).

I instructed coders regarding the use of the coding scheme and monitored their progress (as discussed in section 4.8.3). When all coding was completed, I collated the data and analysed it by using a range of statistical techniques to test the hypotheses of the present study (see section 4.9 for an introduction of the approach to data analysis).

4.8.2 Coder Qualifications

Krippendorff (2004, p. 127) suggested three requirements for coder qualification: cognitive abilities, background, and frequency.

Cognitive abilities. Coders have to understand the procedures involved in conducting content analysis and consistently follow those procedures. In this study, all coders were either postgraduate students or university graduates. Therefore, they had sufficient

cognitive abilities to understand the meaning of the coding scheme and to deal with large quantities of data in a consistent way.

Background. Coders need to have prior experience in contexts similar to the context of the study. In this study, coders were selected to ensure that they were familiar with the content of the constructs for which they coded. For dimensions of knowledge construction, I selected as coders PhD students with teaching experience, and thus with the experience of marking student work to recognize higher level thinking (Coder 1 and Coder 2 in Table 4-3). For dimensions of perceived information quality, I also selected coders who were PhD students, accustomed to consuming large quantities of information (Coder 3 and Coder 4). For information integrity, I used subject matter experts as expert coders (Expert coder 1 and expert coder 2 in Table 4-3)—health practitioners with nursing degrees. Two separate coders, university graduates (Coder 5 and Coder 6), extracted the recommendations from the text before they were rated by health practitioners.

Although Krippendorff (2004) does not recommend the researchers to act as coders, it is a common practice (for content analysis studies of online discussions, this aspect was discussed in detail in section 2.6.6.2). In the present study, to keep research expenses within budget, I acted as a coder in coding the dimensions of knowledge construction (I fulfilled the requirement of being a PhD student with teaching experience). The risk of self-applying coding instructions was managed by executing reliability checks against codes for the same constructs generated by an independent coder who was not a researcher involved in this study.

Table 4-3 *Coders*

Coder	Background	Constructs
1 2[a]	PhD students with teaching experience	Dimensions of knowledge construction: explicitation, evaluation
3 4	PhD students	Dimensions of perceived information quality: relevance, understandability, usefulness
5[b] 6[b]	University graduates	
Expert coder 1 Expert coder 2	Health practitioners with nursing degrees	Information integrity

[a]I, myself, acted as coder 2. [b]Coders 5 and 6 acted as unitizing coders; they extracted recommendations to be rated by the expert coders for information integrity.

Frequency. Coders should be chosen from populations that are accessible by other researchers attempting to replicate the study. Most researchers can easily employ PhD students (including PhD students with teaching experience). Health practitioners with nursing degrees are also relatively easy to access.

4.8.3 Coding Scheme Development, Coder Training, and Coding Procedures

Krippendorff (2004) recommended that coding instructions should be pretested to ensure that formal coding rules are easy to understand and result in high enough inter-rater reliability (and thus, are likely to be valid, in terms of reflecting the objective values of analytic constructs rather than the subjective perceptions and interpretations of the coders).

In the present study, the coding instructions for knowledge construction were pretested. To perceived information quality the study reused a validated research instrument from prior research (McKinney et al., 2002); hence, I believed that pretesting it was less important. For information integrity, a novel measurement instrument was developed; nonetheless, it was not feasible to pretest it, because of the need to use coders with suitable medical qualifications (these were available but expensive to employ). Overall, the study employed six regular coders and two subject matter experts to ensure that

analytical constructs hypothesized to be related were measured independently and thus to minimize common method bias (Malhotra, Kim, & Patil, 2006; Podsakoff et al., 2003). Coders and subject matter experts were employed as research assistants. More intensive pretesting was desirable but not feasible because of financial constraints.

For coding, the following procedure was followed. I met each coder separately (coders never interacted with each other and worked solely on the basis of the coding instructions). I gave each coder the coding instructions and the transcripts to code. Coders were encouraged to ask questions to clarify the task. This initial meeting took about 20 minutes. A second meeting was arranged after the coder completed coding the first 10 threads. I reviewed the results, and the coder had an opportunity to ask further questions. The second meeting took about five minutes. The second meeting allowed me to assess if the coder, indeed, had the cognitive abilities to follow the instructions and to make reasonable judgments. All coders had sufficient cognitive abilities, and there was no need to discontinue the involvement of any of the research assistants based on the results of the second meeting.

For all constructs, there were two independent coders involved to allow reliability checks. All coders coded all of the units; there was no distinction between main and reliability coders in this study.

To test the coding scheme for knowledge construction, a set of 30 scripts not included in the main study was used. The same procedure as in the main study was used, with two coders involved. The coders coded discussion threads for the presence of explicitation. The coders did not encounter problems in using the coding scheme. Based on the outcome of this pilot and to reflect the evolving understanding of the nature of the study, the coding scheme was later updated to use Likert scales to measure the levels of both explicitation and evaluation. To save resources, the new coding scheme was used in the main study without further testing.

4.8.4 Operationalization of Knowledge Construction

Table 4-4 Operationalization of the Dimensions of Knowledge Construction

Construct	Keywords	Indicators
Explicitation		This discussion thread involves cases of
	Refinement	• Weight control related **refinement** (improving or perfecting by pruning or polishing)
	Elaboration	• Weight control related **elaboration** (developing in intricate and painstaking detail)
	Clarification	• Weight control related **clarification** (an interpretation that removes obstacles to understanding)
	Confirmation	• Weight control related **confirmation** (additional proof that something that was believed - some fact or hypothesis or theory - is correct)
	Negation	• Weight control related **negation** (a negative statement; a statement that is a refusal or denial of some other statement)
Evaluation		This discussion thread involves cases of
	Critical discussion	• Weight control related **critical discussion** (discussion characterized by careful evaluation and judgment)
	Argumentation	• Weight control related **argumentation** (a discussion in which reasons are advanced for and against some proposition or proposal)
	Reasoning	• Weight control related **reasoning** (presentation of reasons and arguments)
	Justification	• Weight control related **justification** (the act of defending or explaining or making excuses for by reasoning)

The dimensions of knowledge construction were explicitation and evaluation. I developed the indicators to measure explicitation and evaluation by putting the

keywords presenting the content of these constructs (see Table 3-1 and Table 3-2) in context of the health issue of weight control. The items are listed in Table 4-4.

The indicators were assessed using an eight point semantic differential scale (from *strongly disagree* to *strongly agree*). A scale with an even number of points was chosen to force the coders to make a choice (see also the discussion of this aspect in section 4.4.4).

4.8.5 Operationalization of the Dimensions of Perceived Information Quality

Table 4-5 Operationalization of the Dimensions of Perceived Information Quality

Construct	Keywords	Indicators
Relevance		The information is
	Applicable	• **Applicable** to my interest in weight control
	Related	• **Related** to my interest in weight control
	Pertinent	• **Pertinent** to my interest in weight control
	Relevant in general	• In general, the information is **relevant** to my interest in weight control
Understandability		The information is
	Clear in meaning	• **Clear** in meaning
	Easy to comprehend	• **Easy** to comprehend
	Easy to read	• **Easy** to read
	Understandable in general	• In general, the information is **understandable**
Usefulness		The information is
	Informative	• **Informative** to making decisions regarding approaches to weight control
	Valuable	• **Valuable** to making decisions regarding approaches to weight control
	Useful in general	• In general, the information is **useful**

For perceived information quality, the coders were asked to code from the perspective of an individual interested in weight control. To gain the perspective of such an individual, the coders read the discussion transcripts.

The dimensions of perceived information quality were relevance, understandability, and usefulness. I adapted the indicators from the McKinney et al. (2002) study to fit the context of online discussions in a health support group devoted to weight loss. The items are listed in Table 4-5.

Following McKinney et al. (2002), the indicators were assessed using an 11 point semantic differential scale (from *strongly disagree* to *strongly agree*).

4.8.6 Operationalization of Information Integrity

As discussed in section 4.6, individual recommendations for health-related behaviour or for health-related interventions were coded for information integrity. (The approach used to extract the recommendations from the discussion transcripts is introduced in section 4.8.7).

To operationalize information integrity, I adopted an approach inspired by the Delphi method (Hasson, Keeney, & McKenna, 2000; Yousuf, 2007). Implementing the full Delphi method was not feasible because of limited resources.

Delphi method involves repetitive assessment of recommendations by a panel of experts. In each round, each recommendation is rated and the outcomes are anonymized and summarized. In the first round, the experts receive only the recommendations to rate; in the subsequent rounds, the experts receive both the recommendations and the summary of the previous round.

In this study, the process was limited to a single round. The expert coders (see section 4.8.2 for coder background and qualifications) received the recommendations (see Appendix D) and rated them on an eight point Likert scale, from *highly anecdotal* to *highly evidence based*. The experts rated the recommendations based on their knowledge and experience only, and did not consult Cochrane (or similar) database of

systematic reviews (experts have very limited time, and consulting a database was not feasible).

The ratings were used in analyses at the level of a discussion thread and, in post-hoc analyses, at the level of a recommendation. In the analyses at the level of a recommendation, the ratings by expert coders were used directly. In the analyses at the level of a discussion thread, the average of the ratings for all recommendations in the thread was used to operationalize the information integrity for the thread.

4.8.7 Identification of Task Related Threads and Coding Procedures for Information Integrity

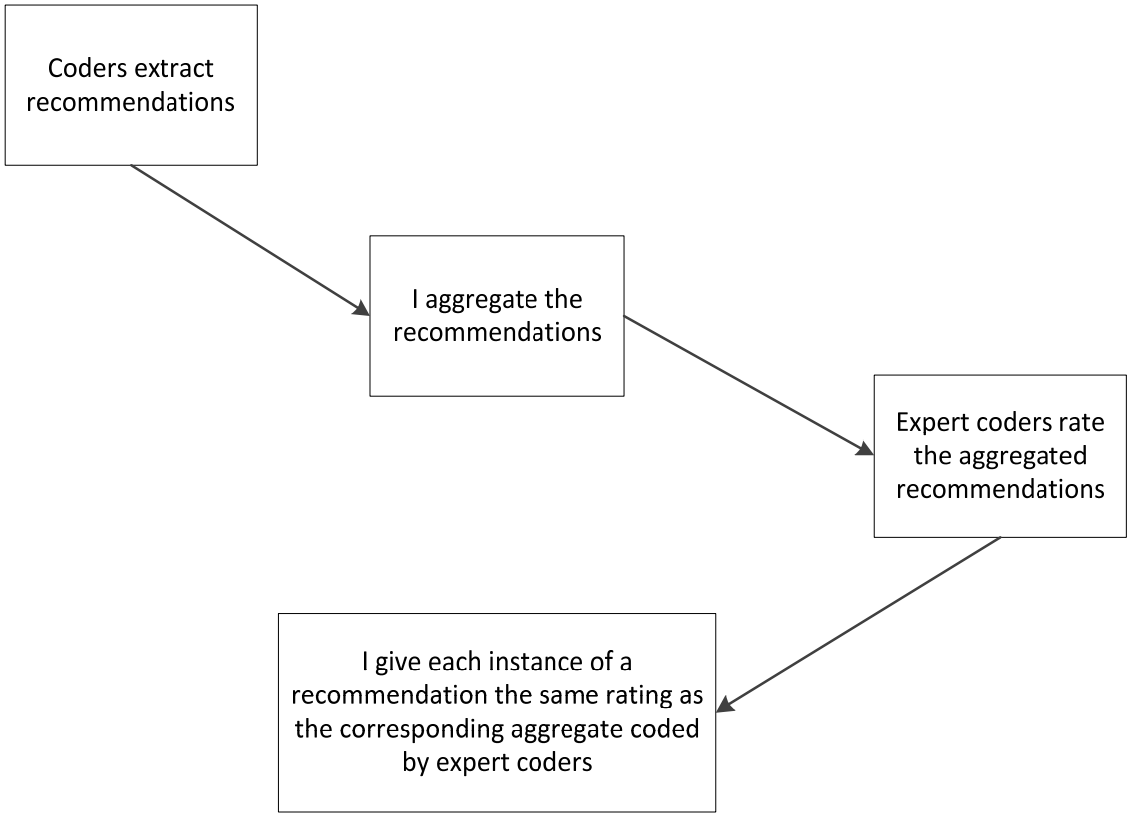

Figure 16. Coding information integrity.

The steps involved in coding information integrity are outlined in Figure 16. Two independent coders (with university education) scanned the discussion for specific recommendations for health-related behaviour or interventions (following the instructions provided in Appendix C). Before passing on the recommendations to

expert coders, I analysed the recommendations and aggregated them to ensure that in the resulting list the same recommendation did not occur more than once (even in a different form). For some of the recommendations, there were multiple instances where a recommendation that was the same in essence was suggested in multiple discussion threads, in a different form. All such instances were aggregated together in the list given to the expert coders and received the same rating. The aggregated recommendations coded by expert coders are listed in Appendix D.

The 108 threads in which both of the coders discovered recommendations were treated as task related; the rest of the threads were not included in further analysis (as described at the end of section 4.7).

4.9 Overall Approach to Data Analysis

Structural equation modelling (SEM), the PLS approach, was used to test the model in Figure 8, with the unit analysis corresponding to a discussion thread. The reason for using SEM rather than linear regression was to assess the discriminant and convergent validity of the latent constructs (the dimensions of knowledge construction and perceived information quality). As an extra check of discriminant and convergent validity, exploratory factor analysis (EFA) was also conducted (with the expectation that the known factor structure would be recovered).

The unit of analysis in SEM (a discussion thread) coincided with the coding unit for the indicators of the dimensions of knowledge construction and perceived information quality. For information integrity, the coding unit was a recommendation; with multiple recommendations possible in a given thread (only threads involving at least one recommendation were included in the analysis, as discussed in section 4.8.7). Information integrity at the level of a discussion thread was operationalized by averaging the information integrity of the recommendations included in the thread, as discussed in section 4.8.6.

As introduced in section 4.8.3, all variables were coded by two coders. The values of variables for EFA and SEM analyses were obtained by adding up the values suggested by the coders. Approaches to assessing inter-rater reliability are discussed in section

4.10.1; the discussion of the reliability of measures of latent constructs in section 4.10.2 assumes that the values suggested by the coders have been added up. In PLS analysis the indicators used to measure constructs were treated as reflective indicators.

4.10 Reliability

A research procedure is reliable if its outcomes for the same phenomena are the same regardless of the details of its implementations (Krippendorff, 2004). When a procedure is unreliable, the chances of obtaining valid results are reduced (but not eliminated altogether). On the other hand, reliability does not guarantee validity, and in the pursuit of high reliability, validity may be lost.

In this study, indicators of latent constructs were estimated by coders coding the transcripts of discussion forum threads. Therefore, two types of reliability were relevant—inter-rater reliability for coding the indicators, and the reliability of the indicators in representing the values of latent constructs. The specific steps used to assess reliability in this study are discussed in the following sections.

4.10.1 Reliability of Coding

Inter-rater reliability was estimated in order to assess the reliability of coding. In this study Likert scales were used, interval scales were assumed, and linear regression and SEM were applied to test the hypotheses of cause-effect relationships. Therefore, consistency between coders in terms of consistently representing changes in values was important, rather consistency at allocating coding units to categories. Therefore, rather than using inter-rater reliability indexes developed to reflect the consistency at allocating units to categories, such as Krippendoff's alpha, this study used Pearson's and Spearman's correlations as inter-rater reliability indexes (as suggested by Riffe et al., 2005).

The present study did not attempt to maximize inter-rater reliability at the expense of validity. In particular, as discussed in section 4.6, rather than using a message or a sentence as a coding unit because smaller units can be coded more reliably, the study used a discussion thread to ensure that the interactions were taken into account in the

measured values of the dimensions of knowledge construction and perceived information quality.

For information integrity the coding unit was not the same as the sampling unit; therefore, the reliability of unitizing was an issue to consider. Nonetheless, as there is no established index for the reliability of unitizing (Krippendorff, 2004, p. 251), this aspect was not addressed.

4.10.2 Reliability of Measures of Latent Constructs

Item reliability. Item reliability refers to how closely the corresponding indicator follows the construct it is supposed to reflect. Factor loadings were used to assess item reliability. Following Nunnally and Bernstein (1994), the cut-off value of 0.7 was used (an item was judged to be reliable if the factor loading was 0.7 or more).

Internal consistency reliability. Internal consistency reliability refers to the extent to which all indicators intended to measure a construct change in a coordinated manner. Internal consistency reliability was assessed by using Cronbach's alpha and composite reliability, with the same threshold of 0.7 (following Nunnally and Bernstein, 1994, and Goodhue, Lewis, and Thompson, 2006). Composite reliability is believed to be a better measure of internal consistency reliability than Cronbach's alpha because it does not assume that all factor loadings are equal; however, to enable comparison with prior studies both of the indexes were assessed.

4.11 Validity

This section discusses the validity of the present study.

4.11.1 Face Validity

Face validity is the extent to which a measure, on inspection, appears to capture the domain suggested by the relevant literature (Krippendorff, 2004). This study reused an existing measure of perceived information quality (from McKinney et al., 2002); therefore, the face validity of the measure was not solely a matter of my own judgment.

For dimensions of knowledge construction, the measure was based on definitions of the categories of explicitation and evaluation by Schellens and Valcke (2005). Thus, again, even though the measure was not reused, its face validity was to an extent supported by prior literature.

4.11.2 Content Validity

Content validity is the extent to which a measure captures the domain of the target construct. Unlike face validity, which is subjective, content validity is an objective property of a measure. Content validity can be tested by formally surveying experts (Lawshe, 1975). An alternative way to ensure content validity is by formulating the measure based on a broad review of the literature.

In this study formal content validity checks (involving a panel of experts commenting on the items included in the measure and on the measure completeness) were not possible because of resource (primarily, time) constraints. Content validity was addressed by relying on a broad review of the literature in formulating the items, and on reusing existing, validated measures whenever possible.

The conceptualization of knowledge construction and the measures of its dimensions was based on a broad review of literature (see section 2.6), and on a synthesis of the literature (see Table 2-18). More specifically, an existing coding scheme was used as a basis, which was selected as the measure best representing the overall body of literature on the topic of knowledge construction in online discussions.

The conceptualization of perceived information quality and the measures of its dimensions were based on the work by McKinney et al. (2002), who based their measure on a broad review of literature.

4.11.3 Nomological Validity

Nomological validity refers to the measure being used to test well-established models and theories and confirming the relationships predicted by these models and theories. This study was to a considerable extent exploratory; hence, the measures I used were

not used in conjunction with strong, well-established theories. Therefore, nomological validity could not be asserted.

4.11.4 Construct, Convergent, and Discriminant Validity

Construct validity refers to whether the measure does in fact measure the construct it is intended to measure. For latent constructs (constructs that cannot be measured directly, as was the case for all constructs employed in this study), construct validity cannot be assessed directly, but face validity (see section 4.11.1), content validity (see section 4.11.2), and nomological validity (see section 4.11.3) contribute to construct validity (that is, face validity, content validity, and nomological validity constitute evidence of construct validity).

Convergent validity and discriminant validity (see section 4.12) are also evidence of construct validity. Convergent validity refers to whether the measure correlates with other (presumed to be valid) measures intended to measure the same construct. Discriminant validity refers to whether the measure is sufficiently different (does not correlate too much) with the measures of other constructs included in the same model. For models involving constructs measured by multiple reflective indicators, there are established statistical procedures that can be used to confirm convergent and discriminant validities of the measures. These procedures are introduced in section 4.12.

4.11.5 Internal Validity

Internal validity refers to whether the nature of the relationships between constructs is as hypothesized in the study. In context of this study, a major aspect of internal validity is whether cause-effect relationships hypothesized in the study refer to actual causes and effects.

This study is a correlational study, and cause-effect relationships cannot be verified directly. One can argue that the nature of this study is such that causes and effects are not likely to be mixed up. Clearly, the transcript is first created (knowledge construction happens), and only then is the resulting content available for assessment in

terms of perceived information quality and information integrity. Nonetheless, one can also argue that in asynchronous discussions contributors read the partial transcript of the discussion as they are involved in the discussion; hence information quality may be affecting the knowledge construction in the discussion.

Experiments can be used to directly and rigorously test the existence of cause-effect relationships. However, as discussed in section 4.3.4, experimental studies of knowledge construction in online discussions are problematic because such discussions are difficult to control.

An alternative approach to clarifying cause–effect relationships is by conducting in-depth qualitative analysis. This is clearly feasible for health support group online discussions, but was beyond the scope of this study.

4.11.6 External Validity (Generalizability)

External validity refers to the extent to which the results of the study can be generalized to a real-world population. This study was limited to a particular period of time and to a particular discussion forum. Therefore, one cannot claim generalizability based on how the content was sampled. On the other hand, as the forum and the content were chosen to be typical of this kind of content, one can claim analytical generalizability (Bedi & Alexander, 2009; Taber, 2000; Yin, 2009)—one can argue that if the hypothesized reasons for the relationships discovered in this study are correct, the results can be generalized to similar settings (where the same reasons apply) that were not part of the population sampled in the research.

4.11.7 Ecological Validity

Ecological validity is the extent to which the study setting approximated a real-world setting. This study used a transcript of a real world-discussion group. Therefore, the ecological validity was high.

4.11.8 Social Validity

Social validity refers to whether or not the research positively contributes to a public discussion of important issues.

This research directly relates to patient centred care and to the danger of Internet users acting on incorrect healthcare related information they obtain from the Internet (the literature highlighting these points was discussed in section 2.3). To ensure that the contribution of this research is positive (and thus to ensure social validity), utmost care was taken to ensure that both the implications and the limitations of this research are well understood and presented explicitly.

4.12 Procedures Used to Assess Convergent and Discriminant Validity

Convergent validity was introduced in section 4.11.4 as the extent to which a measure correlates with other (presumed to be valid) measures intended to measure the same construct. In context of SEM modelling involving latent constructs measured by reflective indicators, it is a common practice to interpret convergent validity as the extent to which the indicators associated with a latent construct behave as if they are representing the values of the same construct.

Item reliability and internal consistency reliability (see section 4.10.2) constitute evidence of convergent validity. In addition, the value of average variance extracted (AVE) (in a SEM model, the average variance predicted in an indicator of a latent variable by that latent variable) for all constructs in the model should be large enough. The commonly used threshold value is 0.5 (Fornell & Larcker, 1981; Henseler, Ringle, & Sinkovics, 2009).

Discriminant validity was introduced in section 4.11.4 as the extent to which a measure is sufficiently different (does not correlate too much) with the measures of other constructs included in the same model. In context of SEM modelling involving latent constructs measured by reflective indicators, it is a common practice to interpret discriminant validity as the extent to which the indicators associated with a latent construct are associated with their own construct more than with other constructs in the model.

A commonly used criterion of discriminant validity is whether each indicator in the model loads on its own construct higher than on other constructs in the model (Hulland, 1999). Another commonly used criterion is whether the square root of AVE for each latent construct is greater than the correlations of the construct with other constructs in the model (Fornell & Larcker, 1981).

An alternative approach to testing convergent and discriminant validity is by applying exploratory factor analysis (EFA) to the indicators of the latent constructs. The EFA procedure is not aware of how indicators are allocated to measures latent constructs, but attempts to discover the factor structure suggested by the data. If the factor structure inductively discovered via EFA coincides with how the latent constructs were operationalized, it constitutes an evidence of both convergent and discriminant validity.

4.13 Approach to Hypotheses Testing

PLS SEM was used to test the hypotheses formulated in section 3.3. A SEM technique was used (rather than linear regression) because SEM explicitly models latent variables, and thus provides information about the fit of the model. PLS was used rather than a covariance based technique (such as the approach used in AMOS software, Arbuckle, 1997) because the study was to a large extent exploratory and because the size of the data used to fit the model was relatively small (Hair, Ringle, & Sarstedt, 2011). By considering sample sizes used in fitting models of similar size using PLS in the literature (Goodhue et al., 2006), it was established that approximately 100 coded threads were needed to test the model. As presented in section 4.7, 108 coded threads were available.

To further confirm the convergent and the discriminant validity, as a supplementary technique, EFA with Maximum Likelihood estimator (Muthen & Muthen, 2007) was also used.

Overall, the analysis involved two stages: testing the measurement model and testing the structural model.

Testing the measurement model. To test the measurement model, a PLS model corresponding to the model in Figure 8 was first fitted to the data. Then, the following checks were conducted:

- item reliability (via item loadings);
- internal consistency reliability (via Cronbach's alpha and composite reliability);
- convergent validity (via AVE values);and
- discriminant validity (by checking for cross-loaded indicators and by comparing square roots of AVE to correlations between constructs).

The details of these checks, including the cut-off values, were presented in sections 4.10 and 4.12.

When problems were discovered (such as cross-loading indicators), they were corrected by removing indicators. When indicators were dropped, the implications of this for the content validity of the measures were carefully considered.

To further confirm the convergent and the discriminant validity, EFA analysis was also conducted to see if the factor structure recovered by EFA matched the structure of the measures known a priori.

Testing the structural model. To test the structural model, an updated PLS model (with indicators dropped when necessary to resolve problems with measures) was fitted to the data. To determine the statistical significance of the path coefficients, a bootstrapping technique was used. As is the common practice in using PLS, the fit of the structural model was judged by the values and the statistical significance of path coefficients, as well as by the values of variance explained R^2 for the dependent constructs. Effect sizes (β) were judged following Kline (1998), based on the values of path coefficients, with values close to or below 0.1, close to 0.3, and close to or above 0.5 judged to correspond to weak, medium, and large effects, respectively.

4.14 Assumptions of Statistical Tests

PLS, unlike covariance-based SEM (Kline, 1998) does not assume multivariate normality (Tenenhaus, Vinzi, Chatelin, & Lauro, 2005). Nonetheless, PLS analysis

achieves greater statistical power when the data are multivariate normal (Marcoulides & Saunders, 2006). EFA analysis using Maximal Likelihood estimator does assume multivariate normality, even though it is common to use EFA with data that are not normal (Muthén & Muthén, 2007). One-way ANOVA analysis and linear regression (used in some of the post-hoc analyses) also assume normality.

To estimate normality, I followed Kline (1998) and assessed the values of skewness and kurtosis for individual variables. According to Kline, a variable is considered to be close to normal if skewness is less than 0.3 and kurtosis is less than 0.8; kurtosis and skewness are indicators of univariate normality for individual variables, which is a prerequisite of multivariate normality. As a direct, stronger test of multivariate normality I also used the Shapiro-Wilk's multivariate normality test.

4.15 Software Programs Used for Statistical Analysis

SPSS (version 18) was used to calculate descriptive statistics and ANOVA.

R (version 2.14.1) was used to estimate and test correlations, to calculate skewness and kurtosis (package moments version 0.13), and to execute the Shapiro-Wilk's multivariate normality tests (package mvnormtest version 0.1-7).

SmartPLS (version 2.0) was used for PLS analysis.

Mplus (version 6.1) was used for exploratory factor analysis.

4.16 Summary

The overall approach to research was positivist, quantitative content analysis. This study was to a large degree exploratory as it was not based on a specific, well-understood theory. Experiment as an approach to hypothesis testing was considered and rejected because of the impossibility to set the levels of knowledge in a discussion. Content analysis was identified as the most suitable approach, because it allowed testing the hypotheses of the study in context of authentic discussions, thus resulting in high ecological validity of the research findings. Nonetheless, the drawbacks of using content analysis were also identified, and included the inability to distinguish causes

from effects by analysing data and the uncertainty of the validity of abductive inferences in relation to the contexts of use.

Then, the approach to content analysis employed in this research was characterized from the perspectives of the criteria and attributes suggested by Krippendorff (2004), including the Krippendorff's criteria for the suitability of content analysis as a technique, the Krippendorff's framework for content analysis, and the Krippendorff's criteria for data languages.

A health support group online discussion site devoted to weight management was the content source in this study. The transcripts were available in the public domain.

The sampling unit was a discussion thread. For the dimensions of knowledge construction (explicitation and evaluation) and for the dimensions of perceived information quality the coding unit was also a discussion thread; for information integrity the coding unit was a recommendation for health-related behaviour or interventions made in a discussion thread. By choosing a discussion thread as a coding unit, the interactions between the contributors were taken into account.

Content was sampled from threads initiated between the first week of August 2010 and the third week of October 2010. Criteria were established for typical threads based on the number of messages and the number of contributors; only typical messages were used in the analysis. As health-related recommendations were identified and extracted, threads with no recommendations were identified and excluded from the analysis as not task related.

Coders from suitable backgrounds were employed as research assistants; to maximize replicability, coder training was not used, and coders could not communicate with each other.

The operationalization of the dimensions of knowledge construction (explicitation and evaluation) was based on the keywords used to describe the concepts of explicitation and evaluation in a prior study. For perceived information quality, the items were adapted from a prior study to fit the context of discussions on the topic of weight management. For information integrity, an approach inspired by the Delphi method was

used, with recommendations extracted from the discussions aggregated and rated by expert coders.

Likert scales were used for coding to enable the use of SEM techniques in data analysis; correspondingly, inter-rater reliability was assessed as correlation, by using Pearson's correlation and Spearman's rho as inter-rater reliability indexes. Item reliability (via factor loadings) and internal consistency reliability (Cronbach's alpha and composite reliability) were used to assess the reliability of the dimensional latent constructs.

The following aspects of validity were considered: face, content, nomological, construct, convergent, discriminant, internal, external, ecological, and social validity. Construct validity was assessed via convergent and discriminant validity. Convergent and discriminant validities were assessed by following the common practice in PLS SEM analysis (via the value of AVE for convergent validity and via cross loadings and by comparing square root of AVE to correlations between latent constructs for discriminant validity). As an alternative assessment of convergent and discriminant validity, EFA was used. Generalizability claimed was analytic generalizability, rather than generalizability to a well-defined population. Ecological validity and social validity were judged to be high because of the study's use of real data and because of the immediate relevance of the study to practice.

PLS SEM (along with bootstrapping) was used to test the structural model, thus testing the research hypotheses.

Even though PLS SEM does not assume normality, it is known to achieve higher statistical power with data distributed close enough to normal. EFA with ML estimator, linear regression, and ANOVA, also used in this study, do assume normality. Normality was assessed by considering skewness and kurtosis and by using the Shapiro-Wilk's multivariate normality test.

Chapter 5. Results and Analysis

5.1 Introduction

This chapter presents the results of data analysis.

Descriptive statistics (minimum, maximum, mean, and standard deviation for each item, separately by coder) and inter-rater reliability (in terms of Pearson's correlation and Spearman's rho) are discussed, followed by the assessment of normality and by the results of a preliminary test of convergent and discriminant validity using exploratory factor analysis.

Then, the results of measurement and structural model testing using PLS SEM are introduced. Even though PLS SEM assesses the measurement and the structural model simultaneously, the chapter follows the common practice and clearly distinguishes the two analyses.

The chapter continues by presenting a summary of a range of post-hoc analyses conducted to further explore the meaning of the results, to address discriminant validity issues discovered in fitting the original model, and to better understand the implications of low inter-rater reliability of some of the measures.

5.2 Descriptive Statistics

Descriptive statistics for the indicators (minimum, maximum, mean, and standard deviation), separately for each coder, are given in Table 5-1, Table 5-2, and Table 5-3.

Table 5-1 *Descriptive Statistics for Knowledge Construction*

	Coder 1				Coder 2			
Indicator[a]	Min	Max	Mean	SD	Min	Max	Mean	SD
Explicitation								
Refinement	0	7	4.73	1.04	2	6	4.49	0.87
Elaboration	0	7	4.57	1.07	2	6	4.41	0.87
Clarification	2	6	4.48	0.94	2	6	4.42	0.82
Confirmation	2	7	4.69	0.92	2	6	4.49	0.83
Negation	0	5	2.78	1.11	1	5	2.55	0.63
Evaluation								
Critical discussion	0	6	3.22	1.46	1	5	3.33	1.02
Argumentation	1	6	3.76	1.47	1	6	3.47	1.15
Reasoning	1	7	3.80	1.49	1	6	3.53	1.16
Justification	0	6	2.96	1.40	1	5	3.19	0.99

[a]For indicator details, refer to section 4.8.4 and Table 4-4.

As seen in Table 5-1, the means for the indicators of explicitation suggest that refinement, elaboration, clarification, and confirmation were roughly equally present, but negation was relatively weak. The means for the indicators of evaluation suggest that critical discussion, argumentation, and reasoning were roughly equally present, but justification was somewhat weaker. All indicators in Table 5-1 were measured on a scale from zero to seven, and for most indicators both coders were close to employing the full range of the values available.

Table 5-2 Descriptive Statistics for Perceived Information Quality

Indicator[a]	Coder 3				Coder 4			
	Min	Max	Mean	SD	Min	Max	Mean	SD
Relevance								
Applicable	2	10	6.66	2.11	0	10	3.55	2.79
Related	1	10	6.71	2.21	0	10	4.38	2.72
Pertinent	1	10	6.61	2.23	0	10	3.94	2.88
Relevant in general	1	10	6.67	2.15	0	10	3.86	2.80
Understandability								
Clear in meaning	4	10	7.54	1.20	3	10	7.52	1.37
Easy to comprehend	4	10	7.66	1.22	4	10	7.77	1.02
Easy to read	3	10	7.68	1.48	3	9	7.34	1.07
Understandable in general	4	10	7.65	1.26	5	9	7.55	0.93
Usefulness								
Informative	1	10	6.05	2.63	0	10	3.90	2.76
Valuable	1	10	6.05	2.93	0	10	3.74	2.98
Useful in general	1	10	6.10	2.81	0	10	3.69	2.92

[a]For indicator details, refer to section 4.8.5 and Table 4-5.

As seen in Table 5-2, indicators of relevance were roughly equally strong; as were the indicators of understandability and usefulness. The mean values of the indicators of relevance and usefulness differed considerably between the two coders; this does not necessarily suggest lack of reliability—because of the way data were analysed in this study, correlations rather than exact matches mattered for the analysis (this aspect is discussed in more detail in section 4.10.1). Inter-rater reliability results in terms of correlations are discussed in section 5.3. All indicators in Table 5-2 were measured on a scale from zero to 10, and for most indicators both coders were close to employing the full range of the values available.

Table 5-3 *Descriptive Statistics for Information Integrity*

	Expert coder 1				Expert coder 2			
Indicator	Min	Max	Mean	SD	Min	Max	Mean	SD
Average	1	7	4.39	1.76	0	7	3.96	1.44

As discussed in section 4.8.6, recommendations extracted from discussion threads, rather than threads themselves were rated for information integrity; more than one recommendation could be present in a discussion thread (but threads with no recommendations were not admitted for analysis; therefore, there was at least one recommendation in each thread). To conduct analysis at the level of a discussion threads, the average of information integrity for each thread was employed as the sole indicator of information integrity. Information integrity was measured on a scale from zero to seven, and both coders were close to employing the full range of the values available.

5.3 Inter-Rater Reliability

As discussed in section 4.10.1, inter-rater reliability was assessed by using Spearman's rho and Pearson's correlation as reliability indices. To judge the levels of reliability, the following thresholds were considered. First, Straub, Boudreau, and Gefen (2004), based on a review of methodological literature, suggested the threshold of 0.7 for both Spearman's rho and Pearson's correlation to assess inter-rater reliability. Second, Cohen (1992) suggested interpreting the size of effects assessed via Pearson's correlation as follows: values close to 0.1 or below, close to 0.3, and close to or above 0.5 correspond to small, medium, and large (or strong) effects, respectively (henceforth, I refer to the Cohen's effect size as "small / medium / strong effect according to Cohen"). Finally, the statistical significance of correlations has also been used to assess inter-rater reliability (Gebauer, Tang, & Baimai, 2008).

As seen in Table 5-4, most of the indicators of evaluation and about half of the indicators of explicitation were above the threshold of 0.7 for inter-rater reliability. For all of the indicators of the dimensions of knowledge construction the Pearson's

correlation corresponded to a strong effect according to Cohen and all of the correlations were statistically significant at $p<0.001$.

Table 5-4 Inter-Rater Reliability for the Indicators of the Dimensions of Knowledge Construction

Indicator	Spearman's rho[a]	Pearson's correlation[a]
Knowledge Construction - Explicitation		
Refinement	0.757	0.727
Elaboration	0.744	0.715
Clarification	0.682	0.632
Confirmation	0.638	0.592
Negation	0.539	0.537
Knowledge Construction – Evaluation		
Critical discussion	0.773	0.759
Argumentation	0.758	0.752
Reasoning	0.756	0.741
Justification	0.639	0.654

[a]For all of the indicators, both of the correlations were statistically significant with $p < 0.001$.

As seen in Table 5-5, none of the indicators of relevance or usefulness were above the threshold of 0.7. Nonetheless, for all of the indicators of relevance and usefulness the values of Pearson's correlations corresponded to strong effects according to Cohen, and all of the correlations were statistically significant at $p<0.001$.

Even so, the inter-rater reliability for understandability was poor. The strength of effect according to Cohen was weak to medium, and for only one of the indicators correlation was marginally statistically significant at the weak alpha protection level of 0.05 (for "easy to read", with $p=0.04$ for Spearman's rho and $p=0.06$ for Pearson's correlation). As the same coders coded all dimensions of perceived usefulness, and for relevance and usefulness the inter-rater reliability was relatively high, the low inter-rater reliability for understandability was, most likely, not because the coders were negligent, but because of the coding scheme or because of the nature of the construct of understandability. The

resources necessary to recode the content with an updated coding scheme were not available; therefore, this issue was not investigated further.

Table 5-5 Inter-Rater Reliability for the Indicators of the Dimensions of Perceived Information Quality

Indicator	Spearman's rho[a]	Pearson's correlation[a]
Perceived information quality – Relevance		
Applicable	0.530 (<0.001)	0.512 (<0.001)
Related	0.531 (<0.001)	0.525 (<0.001)
Pertinent	0.547 (<0.001)	0.518 (<0.001)
Relevant in general	0.545 (<0.001)	0.522 (<0.001)
Perceived information quality - Understandability		
Clear in meaning	0.122 (0.210)	0.113 (0.245)
Easy to comprehend	0.161 (0.097)	0.124 (0.202)
Easy to read	0.198 (0.040)	0.183 (0.059)
Understandable in general	0.163 (0.093)	0.142 (0.143)
Perceived information quality - Usefulness		
Informative	0.547 (<0.001)	0.531 (<0.001)
Valuable	0.614 (<0.001)	0.599 (<0.001)
Useful in general	0.622 (<0.001)	0.614 (<0.001)

[a]Correlation estimate is followed by the *p* value in parentheses.

As seen in Table 5-6, the inter-rater reliability for information integrity was relatively weak, with correlation values well below the threshold of 0.7, but with the value of Pearson's correlation corresponding to a medium to strong effect according to Cohen. Still, both correlations were statistically significant at $p<0.001$.

Table 5-6 *Inter-Rater Reliability for Information Integrity*

Indicator	Spearman's rho[a]	Pearson's correlation[a]
Average	0.431 (<0.001)	0.419 (<0.001)

[a]Correlation estimate is followed by the *p* value in parentheses.

The inter-rater reliability obtained in this study should be considered in context of other studies of knowledge construction (see section 2.6.6.2). Even though it is relatively easy to obtain high inter-rater reliability for variables trivial in meaning (Krippendorff, 2004), the constructs considered in this study are complex. At the same time, none of the studies discussed in section 2.6.6.2 that employed independent coder achieved consistently high inter-rater reliability rates, even though these studies used a message as a unit of analysis (and thus, used a unit of analysis smaller and less complex than a discussion thread).

Based on the results for inter-rater reliability presented in this section, the understandability dimension of perceived information quality was removed from the research model. For the remaining constructs, all indicators were retained, and the values suggested by the coders were added up for use in further analysis.

5.4 Normality

Approaches used in this study to assess normality were introduced in section 4.14. Kurtosis and skewness were used as normality indicators; the Shapiro-Wilk's statistical test for multivariate normality was also applied.

Table 5-7 presents kurtosis and skewness values for all indicators included in the analysis. Kline (1998) suggested that a variable with kurtosis less in absolute value than eight and skewness less in absolute value than three can be treated as normal in path analysis and co-variance-based CFA and SEM modelling using the ML estimator. As seen in the table, kurtosis values for the indicators in this study were in the range from -1.192 to 1.847, and skewness values—in the range from -1.016 to 0.174. Thus, the data were close enough to normal according to the criteria suggested by Kline.

Nonetheless, the Shapiro-Wilk's multivariate normality test indicated that the data were not multivariate normal (p<0.001).

Table 5-7 *Kurtosis and Skewness*

Item	Kurtosis	Skewness
Explicitation		
Refinement	1.847	-1.016
Elaboration	1.226	-0.735
Clarification	-0.132	-0.667
Confirmation	0.507	-0.874
Negation	-0.413	-0.172
Evaluation		
Critical Discussion	-0.737	-0.064
Argumentation	-1.026	-0.062
Reasoning	-0.992	-0.081
Justification	-0.755	0.174
Relevance		
Applicable	-0.879	-0.060
Related	-0.859	-0.200
Pertinent	-0.911	-0.123
Relevant in general	-0.966	-0.141
Usefulness		
Informative	-1.086	0.037
Valuable	-1.192	0.006
Useful in general	-1.133	0.023
Information integrity	-0.226	0.437

In the following analysis, I relied upon the common practice of using statistical methods requiring normality (including EFA with the ML estimator and path analysis with the ML estimator) even though the data were not normal; I considered the data to be close enough to normal because they fulfilled the criteria suggested by Kline (1998).

5.5 Exploratory Factor Analysis

As introduced in section 4.12, for a preliminary analysis of convergent and discriminant validity EFA analysis was applied. The factor structure recovered via EFA was compared to the factor structure suggested by the allocation of indicators to measures (see sections 4.8.4 and 4.8.5). Information integrity was not included in the analysis because it had a single indicator and thus was not a true latent variable. Geomin rotation (an oblique rotation) was used.

First, EFA analysis involving all indicators of explicitation, evaluation, relevance, and usefulness was conducted. Three of the Eigen values were greater than one (see Figure 17). The scree test was inconclusive: the last substantial drop in the magnitude of the Eigen values was not clearly apparent. Thus, on balance, this suggested that there were three factors (rather than four, corresponding to explicitation, evaluation, relevance, and usefulness).

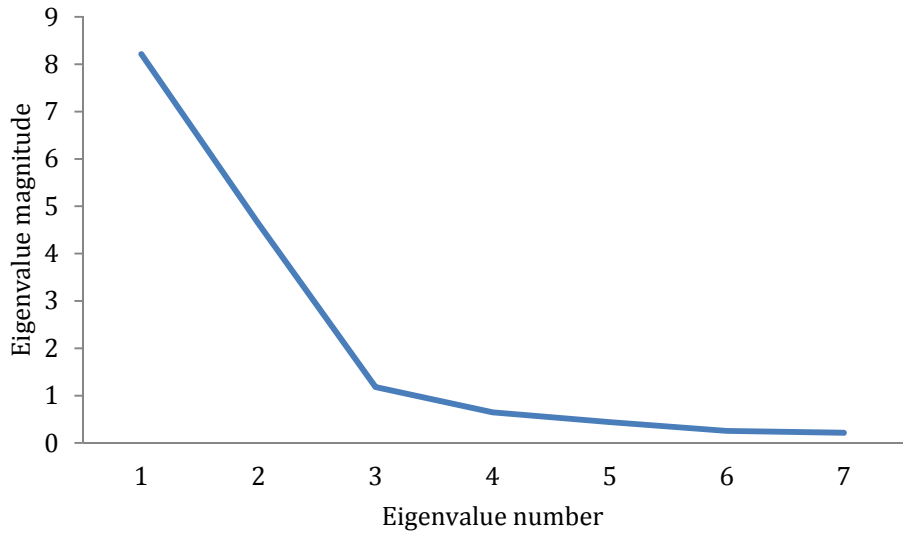

Figure 17. Scree test plot for the indicators of explicitation, evaluation, relevance, and usefulness.

EFA conducted with the number of factors set to four resulted in most of the indicators of explicitation loading highly on Factor 1, apart from negation, which clearly loaded on Factor 2 along with all of the indicators of evaluation (see Table 5-8). Clarification also loaded on Factor 2, marginally higher than on Factor 1. Thus, Factor 1 appeared to

correspond to explicitation, and Factor 2—to evaluation, with the discrimination between the two somewhat weak (unless one accepts that negation should belong to evaluation).

Table 5-8 Factor Loadings in EFA for all Indicators of Explicitation, Evaluation, Relevance, and Usefulness, with the Number of Factors Set to Four

	Factor loadings			
Indicator	Factor 1	Factor 2	Factor 3	Factor 4
Explicitation				
Refinement	0.936	0.025	0.035	-0.029
Elaboration	0.993	-0.027	-0.033	-0.035
Clarification	0.378	0.435	0.001	0.096
Confirmation	0.526	0.308	0.015	0.069
Negation	0.063	0.541	0.034	0.147
Evaluation				
Critical discussion	-0.003	0.953	0.000	-0.014
Argumentation	-0.02	1.013	-0.026	-0.012
Reasoning	-0.007	1.005	-0.009	-0.024
Justification	0.036	0.797	0.039	-0.038
Relevance				
Applicable	0.008	-0.017	0.997	-0.007
Related	0.025	0.005	0.998	-0.035
Pertinent	-0.014	0.017	0.992	0.002
Relevant in general	0.001	0.015	1.005	-0.024
Usefulness				
Informative	0.003	-0.01	0.744	0.393
Valuable	0.002	0.001	0.785	0.364
Useful in general	-0.004	0.009	0.774	0.386

Relevance and usefulness clearly loaded the best on the same factor, Factor 3. Thus, Factor 3 appeared to correspond to perceived information quality, with the two

dimensions of perceived information quality not distinguished. It was not possible to assign meaning to Factor 4.

Thus, the expected structure (four factors corresponding to explicitation, evaluation, relevance, and usefulness) was not recovered. This, however, could be because of a combination of high variable to factor ratio with relatively small number of data points (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

Table 5-9 Factor Loadings in EFA for all Indicators of Explicitation, Evaluation, Relevance, and Usefulness, with the Number of Factors Set to Two

Indicator	Factor loading	
	Factor 1	Factor 2
Explicitation		
Refinement	0.539	0.152
Elaboration	0.519	0.091
Clarification	0.641	0.096
Confirmation	0.596	0.117
Negation	0.573	0.105
Evaluation		
Critical discussion	0.951	-0.004
Argumentation	1.001	-0.032
Reasoning	1.001	-0.019
Justification	0.817	0.031
Relevance		
Applicable	-0.015	0.996
Related	0.016	0.988
Pertinent	0.007	0.991
Relevant in general	0.013	0.995
Usefulness		
Informative	-0.017	0.907
Valuable	-0.006	0.936
Useful in general	-0.002	0.933

EFA for the same set of indicators and with the number of factors set to three resulted in qualitatively the same outcome, with explicitation, evaluation, and the indicators of the dimensions of perceived information quality mostly loading on different factors, very good discrimination between the dimensions of knowledge construction and perceived information quality, and somewhat weak discrimination between the two dimensions of knowledge construction, explicitation and evaluation.

EFA for the same set of indicators and with the number of factors set to two, as expected, resulted in a very clear discrimination between the indicators of the dimensions of knowledge construction and the indicators of the dimensions of perceived information quality (see Table 5-9).

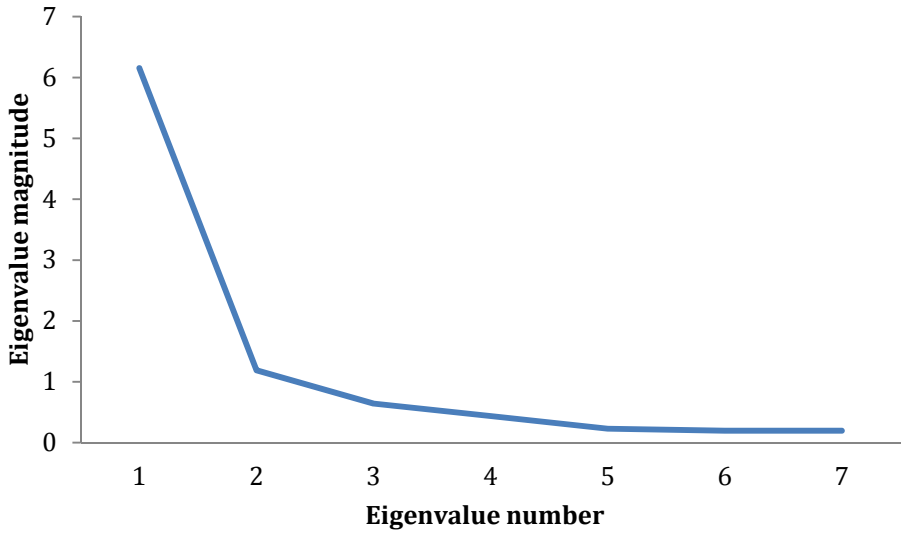

Figure 18. Scree test plot for the indicators of the dimensions of knowledge construction: explicitation and evaluation.

EFA analysis involving only the indicators of knowledge construction, explicitation and evaluation was conducted to see if a better discrimination between explicitation and evaluation could be achieved in an analysis with a higher number of data points per item. Two of the eigenvalues were greater than one; the scree test (see Figure 18) was somewhat inconclusive but consistent with a view that there were two factors.

Table 5-10 Factor Loadings in EFA for Indicators of Explicitation and Evaluation, with the Number of Factors Set to Two

Indicator	Factor loading	
	Factor 1	Factor 2
Explicitation		
Refinement	0.947	0.016
Elaboration	0.984	-0.04
Clarification	0.399	0.426
Confirmation	0.547	0.299
Negation	0.095	0.539
Evaluation		
Critical discussion	0.004	0.948
Argumentation	-0.018	1.005
Reasoning	-0.003	0.999
Justification	0.046	0.796

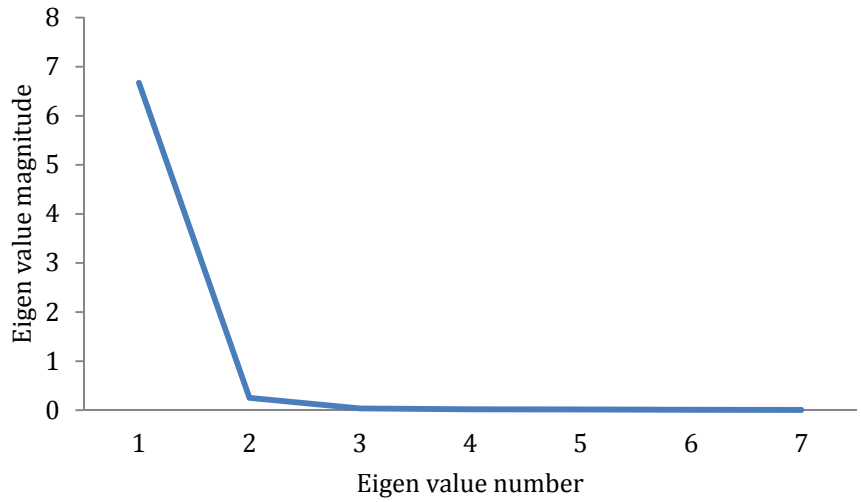

Figure 19. Scree test plot for the indicators of the dimensions of perceived information quality: relevance and usefulness.

EFA conducted with the number of factors set to two resulted in a pattern similar to the pattern in the analysis involving all indicators and with the number of factors set to four

or three (see the discussion earlier in this section): most of the indicators of explicitation loaded highly on Factor 1, apart from negation, which clearly loaded on Factor 2 along with all of the indicators of evaluation (see Table 5-10). Clarification also loaded on Factor 2, marginally higher than on Factor 1. Thus, Factor 1 appeared to correspond to explicitation, and Factor 2—to evaluation, with the discrimination between the two somewhat weak.

EFA analysis involving only the indicators of perceived information quality, relevance and usefulness, was also conducted. Only one of the eigenvalues was greater than one, but the scree test (see Figure 19) suggested quite strongly that there were two factors.

EFA for the indicators of perceived information quality with the number of factors set to two resulted in a clear discrimination between relevance and usefulness (see Table 5-11).

Table 5-11 Factor Loadings in EFA for Indicators of Relevance and Usefulness, with the Number of Factors Set to Two

	Factor loading	
Indicator	Factor 1	Factor 2
Relevance		
Applicable	0.96	0.035
Related	1.021	-0.031
Pertinent	0.942	0.055
Relevant in general	1.005	-0.007
Usefulness		
Informative	-0.04	1.01
Valuable	0.056	0.941
Useful in general	0.003	0.996

Overall, the results of the EFA analysis were inconclusive. In particular, the analysis indicated possible discriminant validity issues between explicitation and evaluation, with one indicator of explicitation, negation, consistently loading on evaluation rather than on explicitation, and another indicator of explicitation, clarification, consistently loading roughly equally high on both of the dimensions of knowledge construction.

Nonetheless, the indicators of the dimensions of knowledge construction and the indicators of the dimensions of perceived information quality did not cross-load; there was a clear discrimination between knowledge construction and information quality.

5.6 PLS Measurement Model

This section discusses the assessment of the PLS measurement model. The procedures used in this study to assess the measurement model were briefly outlined in section 4.13. Reliability of the measures of latent variables was assessed using the approach introduced in section 4.10.2, which involved the assessment of item reliability and internal consistency reliability. Then, convergent and discriminant validity were assessed using the approach introduced in section 4.12. In the rest of this section, these procedures and their outcomes are described in detail.

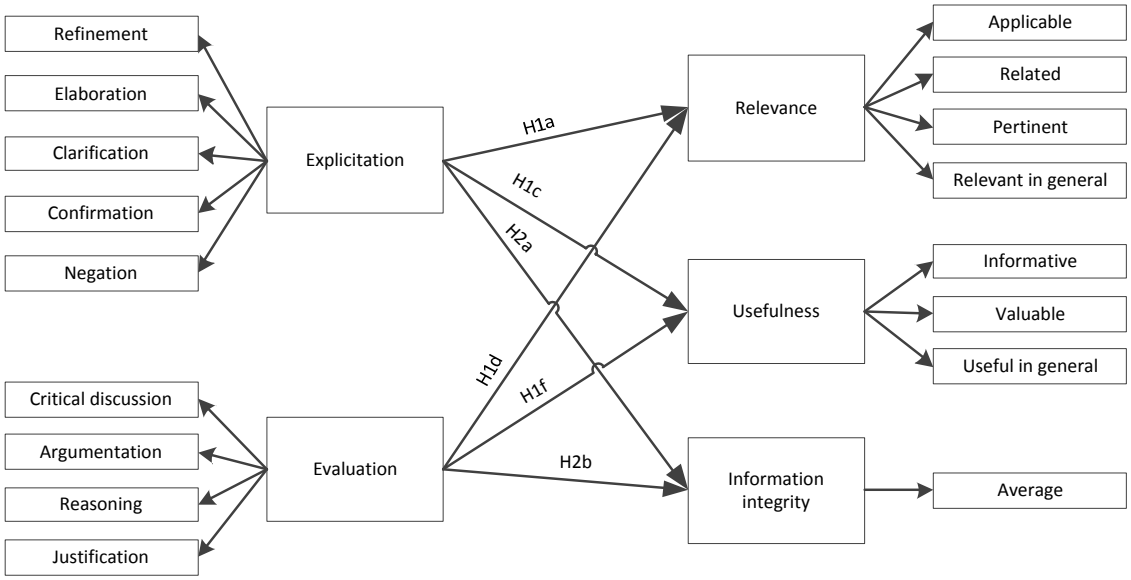

Figure 20. Structural model for PLS analysis.

The model in Figure 8 (with the understandability concept removed, as discussed at the end of section 5.3) is reproduced in Figure 20, with the indicators shown explicitly next to their constructs. The model in Figure 20 was implemented in SmartPLS software, and the PLS algorithm was executed against the data.

Table 5-12 Indicator Loadings on Dimensions of Knowledge Construction

	Explicitation	Evaluation	Relevance	Usefulness	Information integrity
Explicitation					
Refinement	0.862	0.562	0.278	0.259	-0.034
Elaboration	0.846	0.534	0.213	0.197	0.038
Clarification	0.864	0.667	0.246	0.264	-0.002
Confirmation	0.893	0.626	0.255	0.267	-0.080
Negation	0.692	0.619	0.239	0.259	0.114
Evaluation					
Critical discussion	0.704	0.968	0219	0.194	0.055
Argumentation	0.706	0.978	0.202	0.179	0.078
Reasoning	0.714	0.982	0.216	0.187	0.072
Justification	0.655	0.900	0.226	0.184	0.012

Table 5-13 Indicator Loadings on Dimensions of Perceived Information Quality

	Explicitation	Evaluation	Relevance	Usefulness	Information integrity
Relevance					
Applicable	0.289	0.204	0.995	0.924	-0.051
Pertinent	0.293	0.227	0.994	0.927	-0.066
Related	0.305	0.235	0.995	0.914	-0.030
Relevant in general	0.299	0.231	0.998	0.925	-0.054
Usefulness					
Informative	0.287	0.179	0.899	0.987	-0.128
Valuable	0.305	0.197	0.931	0.994	-0.117
Useful in general	0.305	0.202	0.929	0.996	-0.093

Item reliability. To assess item reliability, the loadings of items on their own constructs were examined (see Table 5-12 and Table 5-13). All factor loadings were above the threshold value of 0.7; the only exception was the negation indicator of explicitation, which was very close to 0.7. Thus, in terms of item reliability, the measures were reliable.

Table 5-14 *Indicator Loadings on Information Integrity*

	Explicitation	Evaluation	Relevance	Usefulness	Information Integrity
Average	0.005	0.057	-0.050	-0.113	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Internal consistency reliability. To assess internal consistency reliability, the values of Cronbach's α and of composite reliability (ρ) were examined (see Table 5-15). All of the values were considerably higher than the threshold value of 0.7. (For information integrity, both α and ρ assumed the trivial value of one because it had only one indicator; the notion of internal consistency did not apply.) Thus, in terms of internal consistency reliability, the multi-indicator measures were reliable.

Table 5-15 *Internal Consistency Indices*

	Cronbach's α	Composite reliability (ρ)
Explicitation	0.889	0.919
Evaluation	0.969	0.978
Relevance	0.997	0.998
Usefulness	0.992	0.995
Information integrity	1.000[a]	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Very high levels of the internal consistency indices suggested the possibility that in some cases the meanings of the indicators were too close for the coders to differentiate them clearly. For the dimensions of perceived information usefulness, McKinney et al.

(2002) reported Cronbach α values obtained in a survey setting (with a large number of respondents); the values for relevance and usefulness were 0.85 and 0.88, respectively, which was relatively high, but not as high as in the present study. Because of the difference in research settings (content analysis in this study versus a survey in the study by McKinney et al., 2002), and because the items were reworded for this study, the difference is difficult to interpret.

Convergent validity. Item reliability and internal consistency reliability also constituted evidence of convergent validity. To further explore the convergent validity of the multi-indicator constructs, the values of AVE were examined (see Table 5-16). All of the values were above the threshold of 0.5. Thus, all of the evidence considered suggested that the multi-indicator constructs had convergent validity.

Table 5-16 *Average Variance Extracted*

Construct	AVE
Explicitation	0.696
Evaluation	0.917
Relevance	0.991
Usefulness	0.985
Information Integrity	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Discriminant validity. To assess discriminant validity, loadings of items on their own constructs were compared with their loadings on other constructs in the model (see Table 5-12, Table 5-13, and Table 5-14). For all of the constructs, items loaded on their own constructs higher than on any other constructs in the model, suggesting discriminant validity.

The negation indicator of explicitation also loaded high on evaluation, but it loaded higher on its own construct; this should be contrasted with the EFA analysis results, in which negation loaded considerably higher on evaluation than on its own construct. Even though further investigation of this issue is desirable (indeed, one can argue that negation corresponds to higher level of knowledge construction than explicitation),

larger data sets, and possibly the use of confirmatory covariance-based techniques, are desirable to investigate this aspect further. In this study, I took a view that the PLS result should take precedence. It is common to conduct PLS analysis only, and to accept the PLS results as correct. Moreover, PLS does not assume multivariate normality (unlike EFA with ML estimator), and thus, arguably, fitted better the data of this study (for an assessment of how close the data were to normal, see section 5.4).

In EFA analysis, the clarification indicator of explicitation loaded approximately equally high on the factors corresponding to explicitation and evaluation. In PLS analysis, clarification clearly loaded higher on its own construct. As with the negation indicator of explicitation, the PLS result was given precedence.

The indicators of usefulness also loaded high on the other dimension of perceived information quality, relevance, but they loaded higher on their own construct. Similarly, the indicators of relevance also loaded high on usefulness, but loaded higher on their own construct. This was consistent with the results of the EFA analysis, which suggested that the two dimensions of perceived information quality were close but distinct.

As a further assessment of discriminant validity, the square root of AVE for each latent construct was compared to correlations of the construct with other constructs in the model (see Table 5-17). For all of the constructs, the square root of AVE was larger than the correlations, suggesting discriminant validity. On the other hand, the absolute value of the correlation between explicitation and evaluation was quite high, and the absolute value of the correlation between relevance and usefulness was very high (0.927) putting the discriminant validity in doubt (Hair, 1998, suggested that for discriminant validity correlations between latent variables should not exceed 0.7).

As most of the commonly suggested criteria for discriminant validity were met, I proceeded with the main analysis assuming discriminant validity. Moreover, to further explore the implications of the discriminant validity issues, I conducted a series of post hoc analysis (reported in section 5.8) assuming lack of discriminant validity between the dimensions of knowledge construction (explicitation and evaluation) and between the dimensions of perceived information quality (relevance and usefulness) and treating

knowledge construction and perceived information quality as one-dimensional constructs.

Table 5-17 *Latent Variables Correlation Matrix*

	Explicitation	Evaluation	Relevance	Usefulness	Information Integrity
Explicitation	0.835				
Evaluation	0.726	0.958			
Relevance	0.298	0.225	0.995		
Usefulness	0.302	0.194	0.927	0.992	
Information Integrity	0.005	0.057	-0.050	-0.113	1.000

Note. The values on the diagonal are square roots of AVE.

5.7 PLS Structural Model

This section discusses the results of testing of the PLS structural model. The procedures used in this study to test the structural model were briefly outlined in section 4.13. Bootstrapping with 500 resamples was used to assess the statistical significance of path coefficients. The path coefficients, the corresponding *p* values (obtained via bootstrapping), and the variance explained (R^2) values are shown in Figure 21.

Evaluation was not found to affect any of the dependent constructs, and information integrity was not affected by the dimensions of knowledge construction, and thus was not explained by the model.

Explicitation affected usefulness at alpha protection level 0.01, with medium effect size according to Kline (see section 4.13 for an introduction of how effect sizes were assessed). Furthermore, explicitation affected relevance at alpha protection level 0.05, also with medium size (but the effect was somewhat weaker than the effect on usefulness).

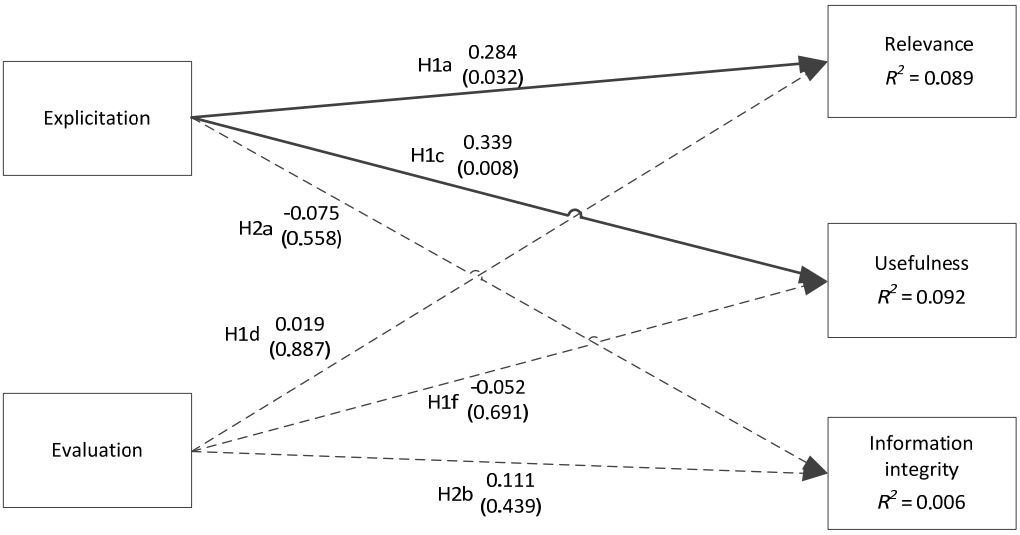

Figure 21. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines.

For both usefulness and relevance, the R^2 values were low, less than 0.1, and hence, "weak" according to the classification by Hair et al. (2011), who suggested the threshold of 0.25 for the amount of variance explained to be considered weak. This outcome is not surprising, and a direct consequence of both usefulness and relevance having only one statistically significant predictor (explicitation) that affected them with a moderate effect size.

5.8 Post-Hoc Analyses

This section presents the results of post-hoc analyses conducted to further clarify the results of measurement and structural model testing presented in sections 5.6 and 5.7.

5.8.1 A Model With Both Knowledge Construction and Perceived Information Quality Treated as One-Dimensional Constructs

In the analysis of the initial research model (see section 5.6) it was found that the indicators of the dimensions of perceived information quality, relevance, and usefulness loaded the highest on their own constructs. It was also found that the square root of

AVE for each of the constructs was higher than the correlation between the constructs. However, the absolute value of the correlation was very high (0.927), which can be seen as a problem with discriminant validity.

The discriminant validity of the dimensions of knowledge construction, explicitation and evaluation was less in doubt. Even though one of the indicators of explicitation, negation, loaded high on both of the constructs, the rest of the indicators clearly loaded the highest on their own constructs. The square root of AVE for each of the constructs was higher than the correlation between the constructs, and the correlation between the constructs was high (0.726), but not as high as between relevance and usefulness. Nonetheless, the correlation was somewhat higher than the threshold of 0.7 suggested by Hair et al. (2009).

Issues around the discriminant validity between relevance and usefulness and between explicitation and evaluation were also indicated in the EFA analysis presented in section 5.5.

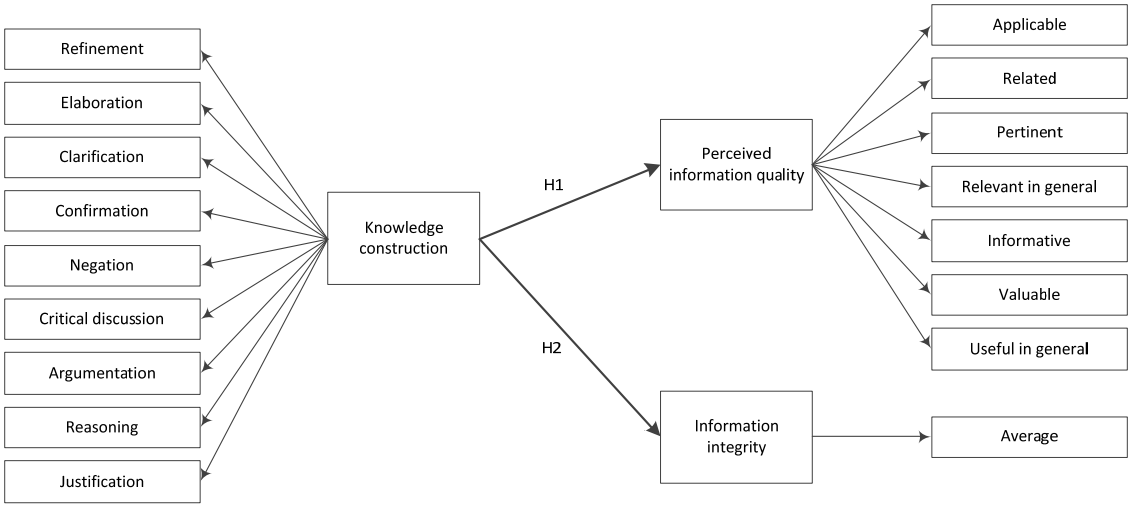

Figure 22. Structural model with both knowledge construction and perceived information quality treated as one-dimensional constructs.

While it is not conclusive that there is no discriminant validity between relevance and usefulness and between explicitation and evaluation, exploring the consequences of viewing the constructs of knowledge construction and of perceived information quality as one-dimensional offers an additional insight for the first of the high-level hypotheses

suggested in this study (H1, see section 1.6). Therefore, in this section I assess a research model with knowledge construction and perceived information quality represented as one-dimensional constructs, with knowledge construction measured by combining all indicators of explicitation and evaluation (from Table 4-4) and perceived information quality—by combining all indicators of relevance and usefulness (from Table 4-5). The model is given in Figure 22. The hypotheses H1 and H2 in the model correspond to the hypotheses introduced in section 1.6.

The details of the measurement model analysis are given in section E.1 of Appendix E. There were no convergent or discriminant validity issues.

The results of structural model testing are presented in Figure 23, including the path coefficients, the corresponding *p* values (obtained via bootstrapping), and the variance explained (R^2) values.

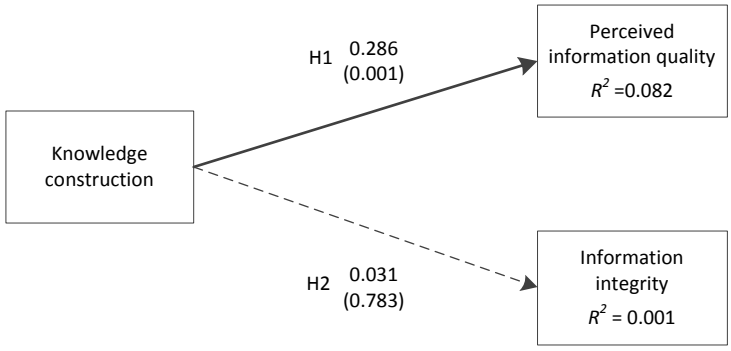

Figure 23. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines.

Information integrity was not affected by knowledge construction, and thus was not explained by the model.

Knowledge construction affected perceived information quality at alpha protection level 0.01, with medium effect size according to Kline (see section 4.13).

The R^2 for perceived information quality was low, less than 0.1; the variance explained was "weak" according to the classification by Hair et al. (2011).

5.8.2 A Model With Explicitation and Evaluation as Separate Constructs and With Perceived Information Quality Treated as a One-Dimensional Construct

In the research model with knowledge construction and perceived information quality treated as multidimensional constructs (see Figure 21), only the explicitation dimension of knowledge construction affected the dimensions of perceived information quality, and, according to the values of the path coefficients, had somewhat higher explanatory power than the one-dimensional construct corresponding to knowledge construction in Figure 23 (in particular, with respect to the usefulness dimension of perceived information quality, with $\beta=0.339$, compared to $\beta=0.286$ for the relationship between knowledge construction and perceived information quality in Figure 23).

As highlighted in section 5.6, the correlation between the dimensions of perceived information quality was very high (0.927) for a model assuming that both knowledge construction and perceived information quality are two-dimensional constructs (see Figure 20). The correlation between the dimensions of knowledge construction was also high (0.726), but considerably lower than for the dimensions of perceived information quality.

In section 5.8.1, a model viewing both knowledge construction and perceived information quality as one-dimensional constructs was considered; although the model did not have any discriminant validity problems, and was more parsimonious than the model in Figure 20, it resulted in a loss of explanatory power, as discussed at the end of section 5.8.1.

In this section, I test a model with knowledge construction treated as a multidimensional construct and perceived information quality treated as a one-dimensional construct measured by combining the items of relevance and usefulness (see Table 4-5).

The model is given in Figure 24. The hypotheses H1ac and H1df in the model aggregate the hypotheses H1a / H1c and H1d / H1f introduced in section 3.3.1. Taken together, the hypotheses H1ac and H1df elaborate the high-level hypothesis H1 in

Figure 2. The hypotheses H2a and H2b are the same as the hypotheses H2a and H2b introduced in section 3.3.2.

The details of the measurement model analysis are given in section E.2 of Appendix E. Even though the correlation between explicitation and evaluation was marginally higher than the threshold value of 0.7 (as in the analysis in section 5.6), there were no other convergent or discriminant validity issues.

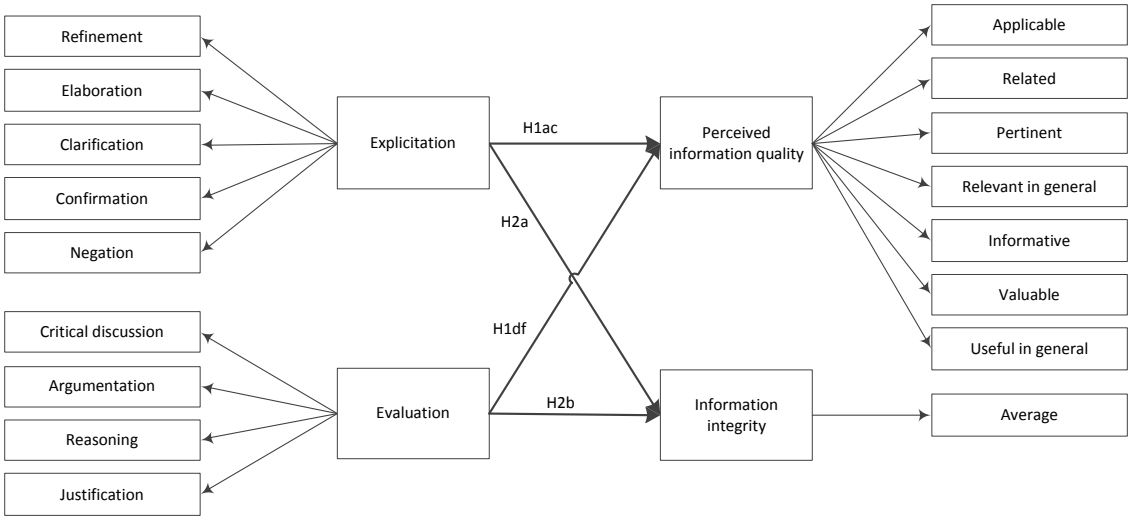

Figure 24. Structural model with perceived information quality treated as a one-dimensional construct.

The results of the structural model testing are given in Figure 25, including the path coefficients, the corresponding *p* values (obtained via bootstrapping), and the variance explained (R^2) values.

Information integrity was not affected by the dimensions of knowledge construction, and thus was not explained by the model.

Explicitation affected perceived information quality at alpha protection level 0.05, with a medium effect size. The effect was stronger than the effect of knowledge construction on perceived information quality in a model with the knowledge construction treated as a one-dimensional construct (see Figure 23).

The R^2 for perceived information quality was low; the variance explained was "weak" according to the classification by Hair et al. (2011). Nonetheless, the R^2 value was

higher than the R^2 value for information quality in a model with both knowledge creation and perceived information quality treated as one-dimensional constructs (0.093 versus 0.082). Arguably, this model provided a better fit to the data comparing to the models considered in sections 5.7 and 5.8.1, as it avoided major discriminant validity issues while retaining explanatory power.

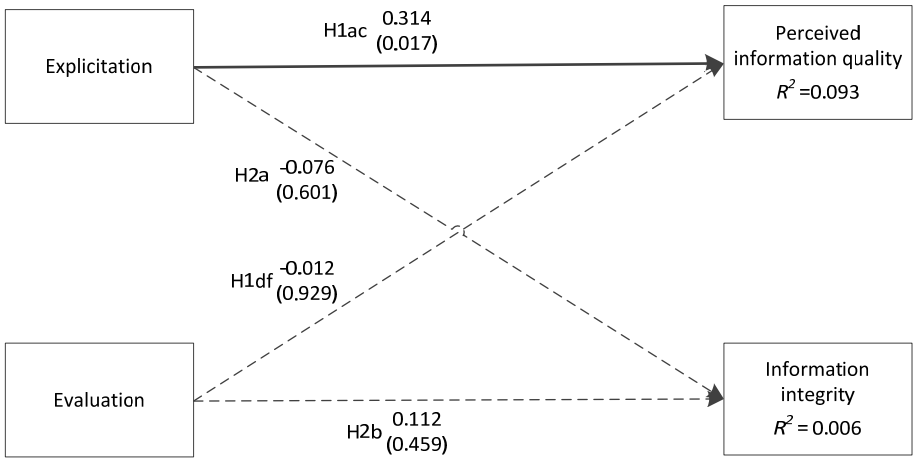

Figure 25. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines.

5.8.3 A Model With Explicitation Hypothesized to Affect Evaluation

It is plausible to suggest that the dimensions of knowledge construction are not independent, and that the activities at lower levels of knowledge construction lead to activities at higher levels of knowledge construction. Addressing this aspect, however, was not the focus of this study. Moreover, the coding procedures were not suitable to explore the relationships between different levels of knowledge construction, because both of the dimensions included (explicitation and evaluation) were coded by the same two coders (therefore, any relationships discovered may reflect the opinions of the coders, rather than the substantive relationships). Nonetheless, exploring the relationship between explicitation and evaluation is clearly of interest.

To explore the relationship between explicitation and evaluation, a model including such a relationship is tested in this section. First, the section provides a justification for

a relationship between explicitation and evaluation. The research model with this hypothesis added is presented in Figure 26. Then, the measurement and the structural model are fitted to the data, following the same pattern as in sections 5.8.1 and 5.8.2.

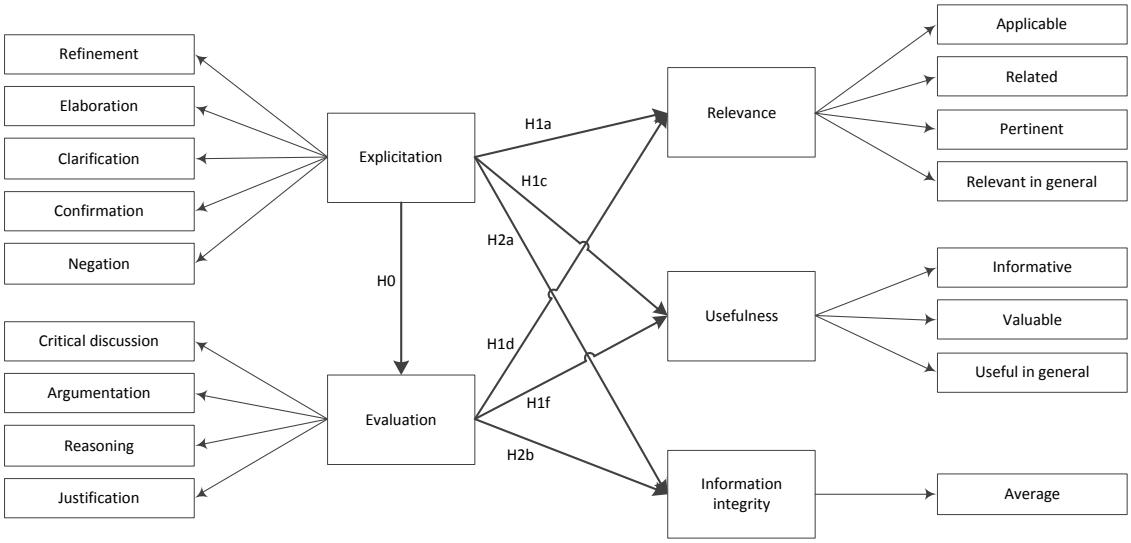

Figure 26. Structural model with the hypothesis that explicitation affects evaluation.

H0: Explicitation → Evaluation. Refinement, clarification, and elaboration of the new information allow the contributors to form the vocabulary that can be used to advance propositions based on reasoning, to justify propositions, to advance arguments in favour of propositions, or to criticize them. In context of health support group discussions, new information regarding various behaviour and interventions, after the initial clarification and refinement, would lead to propositions regarding health-related recommendations that need to be further justified or may be criticized from various perspectives. Therefore, higher levels of explicitation create more opportunities for later evaluation, so that ultimately explicitation causes evaluation. This view is consistent with the view expressed by Schellens and Valcke (2005) that levels of knowledge construction can be seen as phases of knowledge construction, with lower levels enabling higher levels.

The details of the measurement model analysis for the model in Figure 26 are given in section E.3 of Appendix E. Convergent and discriminant validity issues were the same as in the analysis in section 5.6. The results of the structural model testing are given in Figure 27, including the path coefficients, the corresponding *p* values (obtained via bootstrapping), and the variance explained (R^2) values.

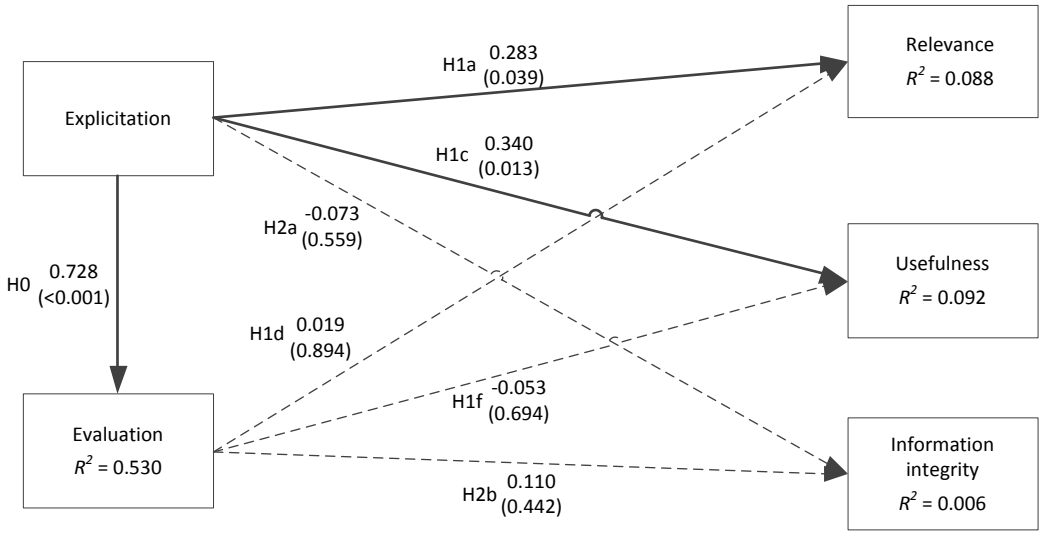

Figure 27. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines.

Information integrity was not affected by the dimensions of knowledge construction, and thus was not explained by the model.

Explicitation affected both dimensions of perceived information quality at alpha protection level 0.05, with a medium effect size.

The R^2 for both dimensions of perceived information quality was low; the variance explained was "weak" according to the classification by Hair et al. (2011).

Explicitation affected evaluation at alpha protection level 0.001, with a strong effect size. The R^2 for evaluation was "moderate" according to the classification by Hair et al. (2011).

Overall, the results were very similar to the results for the model with no relationship hypothesized between explicitation and evaluation. In particular, only evaluation affected the dimensions of perceived information quality, with the effect on usefulness somewhat larger than on relevance.

5.8.4 A Model of Coder's Conceptions

In this section I explore the consequences of different interpretation of coding schemes by coders by explicitly modelling coder conceptions of the constructs. I use hierarchical PLS modelling, as proposed by Wetzels, Odekerken-Schroder, and Van Oppen (2009).

The model is based on an assumption that ratings for items given by each coder represent the coder's conceptions of the corresponding construct formed based on both the coding scheme and on the coder background. Such coder's conceptions are objective in the sense that they are part of the objective (from the perspective of the researcher) reality; they can be measured similarly to other properties of the coder, such as weight, height, or level of intelligence. Moreover, I assume that a coder's ratings of units of coding (in the present study, of discussion threads) reflect the coder's conceptions, which are, in their turn, are affected by the "true" construct intended to be measured by the coding scheme.

The model accounting for coders' conceptions is presented in Figure 28. Constructs labelled with coder numbers (for definitions of coder numbers, see Table 4-3) were measured by the items of the corresponding "true" construct as they were rated by the coder. The "true" constructs were measured as in the rest of the models presented in this thesis, by using a sum of the ratings by both of the coders. Thus, the approach somewhat differs from the one used in hierarchical PLS models introduced in Wetzels et al. (2009), where higher-level constructs were measured by combining all indicators of the lower level constructs (rather than adding them up). The difference was to make the model in this section consistent with other models in this study, to enable comparisons. (It was also believed that combining the indicators, rather than adding them up, may result in problems with statistical analysis when inter-rater reliability is very high, as for "true" constructs the indicators coming from one of the coders would have the same values as the indicators coming from another coder.)

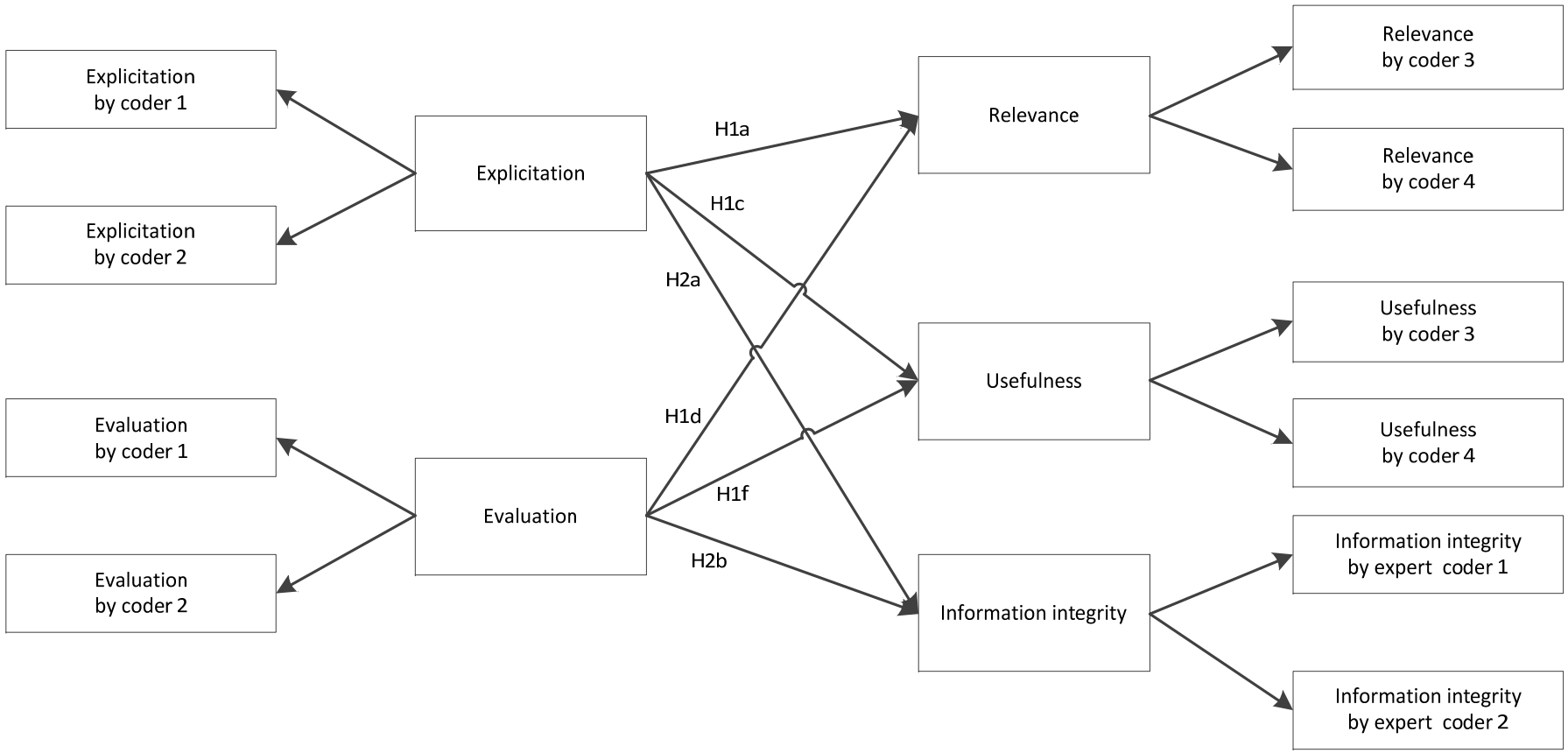

Figure 28. Structural model accounting for coders' conceptions.

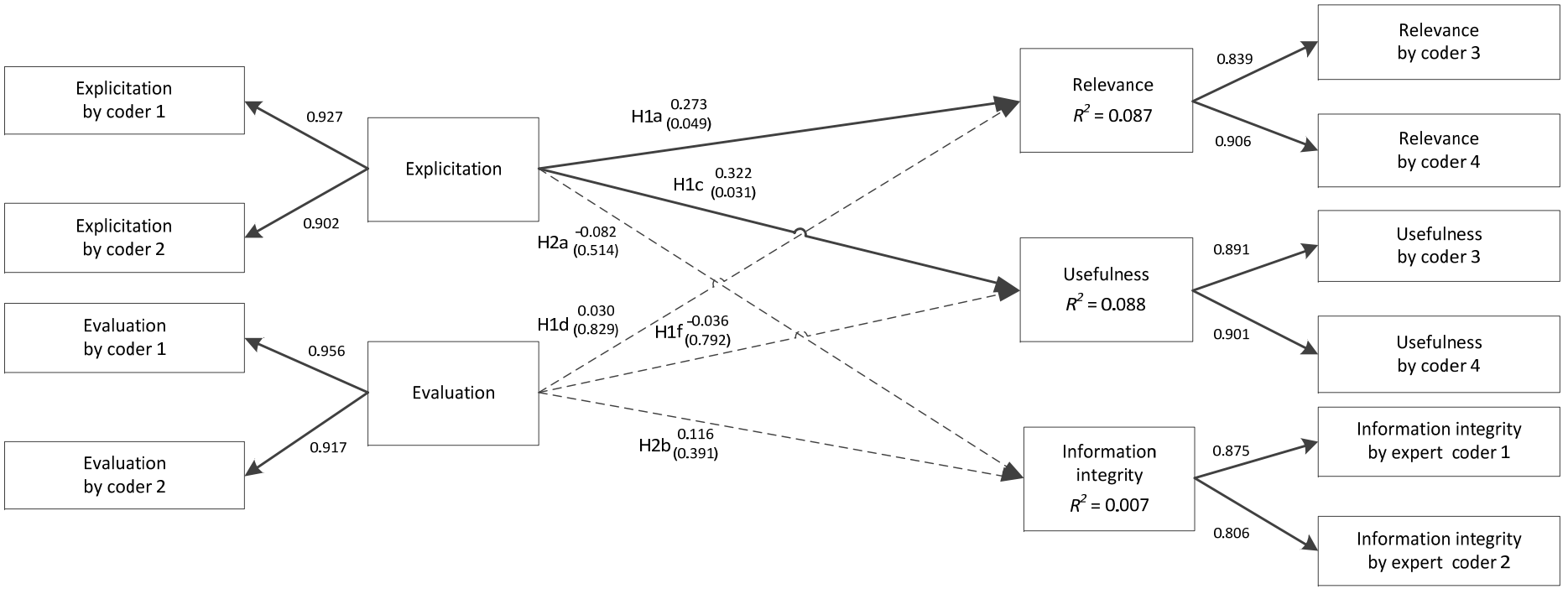

Figure 29. The results of testing the structural model. The values of path coefficients are given next to the corresponding hypotheses labels, with the corresponding *p* values shown in parentheses. For paths between higher-level constructs and their coder's conceptions, all *p* values were below 0.001, and are not shown in the figure. Paths found to be statistically significant at alpha protection level 0.05 are shown as continuous lines; paths not found to be statistically significant are shown as dashed lines.

The details of the measurement model analysis for the model in Figure 28 are given in section E.4 of Appendix E. As expected, there were discriminant validity issues between constructs and their generalizations (immediately related concepts at higher levels in the hierarchical model). In the spirit of the approach to hierarchical PLS modelling suggested by Wetzels et al. (2009), the structural model was fitted without making adjustments to resolve discriminant validity issues between such constructs.

The results of the structural model testing are given in Figure 29, including path coefficients, the corresponding *p* values (obtained via bootstrapping), and the variance explained (R^2) values.

The overall results were very similar to the corresponding model that did not account for coder conceptions (see section 5.7).

The model fulfilled the condition suggested by Chin (1998) that higher-level constructs and the constructs under them should have values for path coefficients at least as high as 0.7. Overall, the results suggested that even though the inter-rater reliability was not high, the coders' concepts were close to the "real" constructs.

5.8.5 Information Integrity: Analysis at the Level of a Recommendation

In the analyses in section 5.7 and in sections 5.8.1 to 5.8.4 the unit of analysis was a thread, and information integrity was operationalized as an average of the information integrity values for all recommendations in the thread. Consistently over all analyses, information integrity was not predicted by the models considered. This, however, may have been because of the choice of the unit of analysis; because information integrity was coded at the unit of a recommendation, analysis at the level of a recommendation was likely to be more sensitive.

An attempt was made to fit a PLS model with explicitation and evaluation hypothesized to affect information integrity at the level of a recommendation, and thus with data involving 180 cases, one case for each recommendation. None of the paths were statistically significant. However, there were problems with convergent and discriminant validity. This was, most likely, because PLS estimates the values of

parameters by maximizing the amount of variance explained; thus, when there is no variance explained, the estimates are not reliable.

Even though the test using PLS indicated that variability in information integrity was not predicted even when the unit of analysis was a recommendation, a similar model was tested using linear regression (using Mplus software). This was to obtain estimates of the *p* value by using a modelling technique that does not break down when there is no variance explained in the model. The same relationships were assumed (information integrity was regressed on explicitation and evaluation); to enable the use of linear regression, explicitation and evaluation were measured by adding up the values of their indicators. For the path from explicitation to information integrity the *p* value was 0.506, and for the path from evaluation to information integrity the *p* value was 0.267.

Thus, no variance in information integrity was explained even when using models at the level of a recommendation.

To explore the possibility that one of the coders coded information integrity correctly, and another—not correctly (resulting in the low level of inter-rater reliability for information integrity), analysis was also conducted by using the rating by each coder, separately. Assuming that only one of the coders rated information integrity correctly, and that there is a relationship either between explicitation or evaluation and information integrity, statistically significant relationships should be discovered for the coder who rated correctly. For coder 1, the *p* values were 0.591 and 0.335 for explicitation and evaluation, respectively. For coder 2, the *p* values were 0.545 and 0.352. Thus, either both of the coders coded incorrectly, or there was no relationship between any of the dimensions of knowledge construction and information integrity.

To explore the issue of variability in information integrity, ANOVA analysis was used to test the hypothesis that in all of the threads information integrity was approximately the same. For information integrity measured by a sum of the coders' ratings, this hypothesis was not rejected (and thus was confirmed) with p=0.327. For the information integrity as rated by code 1 only, the *p* value was 0.373, and for coder 2—0.272. Thus, no matter how knowledge construction is measured, it was unlikely that a relationship between the level of knowledge construction for a thread and information integrity would be found.

5.9 Interpretation of Data Analysis Results

This section discusses and interprets the results of the analyses presented so far in Chapter 5 from two perspectives: from the perspective of conceptualization and operationalization of constructs and from the perspective of the support for the high-level hypotheses stated in Chapter 1 as the research questions addressed in the present study.

5.9.1 Conceptualization of Constructs

The conceptualizations of the constructs of knowledge construction, namely perceived information quality and information integrity, were initially discussed in section 3.2, and the operationalisations—in sections 4.8.4, 4.8.5 , and 4.8.6, respectively. This section discusses the implications of the results of the analysis presented in Chapter 5 for the appropriateness of the initial conceptualizations (and for the operationalisations).

5.9.1.1 Knowledge Construction

In this study, knowledge construction was conceptualized as a multidimensional construct, with the dimensions of explicitation and evaluation.

Inter-rater reliability for the indicators used to measure explicitation and evaluation was relatively high (see Table 5-4 in section 5.3). Three of the four indicators of evaluation and two of the five indicators of explicitation cleared the threshold value of 0.7. The only indicator that had inter-rater reliability below 0.6 was the negation indicator of explicitation.

In an exploratory factor analysis with the number of factors set to four (using the data for the indicators of explicitation, evaluation, relevance, and usefulness), all of the indicators of evaluation loaded the highest on the same factor (see section 5.5). Most of the indicators of explicitation loaded the highest on a separate factor, but the negation indicator of explicitation loaded much higher on evaluation than on explicitation, while the clarification indicator of explicitation loaded high on both of the factors. Thus, overall, the results of the EFA analysis supported the conceptualization of knowledge construction as a multidimensional construct.

In PLS measurement model analysis, models treating knowledge construction as a multidimensional construct (see section 5.6, as well as the measurement model analyses in section 5.8) no major discriminant validity issues between explicitation and evaluation were discovered, although the correlation between the two constructs was relatively high (0.726) and slightly higher than the threshold value of 0.7. Unlike in EFA analysis, the clarification indicator of explicitation did not cross-load. The negation indicator of explicitation also loaded high on evaluation, but loaded the highest on its own construct; the absolute value of the loading of negation on its own construct was relatively low (but only marginally lower than the threshold value of 0.7).

In post hoc analysis, the best model fit (in terms of maintaining the explanatory power while avoiding discriminant validity issues) was achieved in a model with explicitation and evaluation treated as separate constructs.

It should be noted that the approach to coding in this study was more likely to result in bias towards higher correlation between the constructs of evaluation and explicitation because the constructs were coded by the same two coders (because of common method bias) than in bias towards lower correlations between the constructs.

Thus, overall, the results of the analysis supported the conceptualization of knowledge construction as a multidimensional construct. As far as the content of the constructs is concerned, the results suggested that negation may be an indicator of evaluation rather than of explicitation. Indeed, one may argue that negation (and thus, criticism of statements made earlier in a discussion), corresponds to a higher level of learning than clarification (merely clarifying the meaning of statements made earlier in a discussion).

5.9.1.2 Perceived Information Quality

Perceived information quality was initially conceptualized as a multidimensional construct with the dimensions of relevance, understandability, and usefulness.

Inter-rater reliability for all indicators used to measure relevance and usefulness was above 0.5, with two of the indicators of usefulness having inter-rater reliability of very close to or above 0.6 (see Table 5-5 in section 5.3); for all of the indicators of relevance and usefulness the correlations of the ratings by different coders were statistically significant. Even so, inter-rater reliability for the indicators of understandability was

very poor; correlations between the ratings by different coders were not statistically significant. As the result, understandability was removed from further analysis.

As discussed in section 5.3, low inter-rater reliability for understandability was most likely a defect of the coding scheme. The coders were asked to code the dimensions of perceived information quality from the perspective of individuals interested in weight management issues. Thus, lack of reliability reflected the differences between coders in their interpretation of which information is likely to be seen as understandable.

In an exploratory factor analysis with the number of factors set to four (using the data for the indicators of explicitation, evaluation, relevance, and usefulness), all of the indicators of relevance and usefulness loaded the highest on the same factor. Nonetheless, in an exploratory factor analysis with the number of factors set to two, with the data for the indicators of relevance and usefulness only, the indicators of relevance and usefulness loaded on different factors, with very little cross-loading.

In PLS measurement model analysis for models treating perceived information quality as a multidimensional construct (see section 5.6 and section 5.8), correlation between relevance and usefulness was very high (very close to one), suggesting lack of discriminant validity (even though the constructs had discriminant validity according to other criteria).

Overall in post hoc analyses, the best model fit (in terms of maintaining the explanatory power while avoiding discriminant validity issues) was achieved in a model with perceived information quality treated as a one-dimensional construct (measured by combining all indicators of relevance and usefulness).

Even though relevance and usefulness were rated by the same coders, it appears unlikely that the lack of discriminant validity was solely due to common method bias. Most likely, lack of discriminant validity was because the content of the two constructs is very similar. Indeed, from the perspective of an individual interested in weight management, only useful information is truly relevant (and the other way round).

Thus, overall, the results of the analysis did not support the initial conceptualization of relevance and usefulness as separate dimensions of perceived information quality.

Because of the problems with measuring understandability, its conceptualization as a separate dimension of perceived information quality was not tested.

5.9.1.3 Information Integrity

Information integrity was conceptualized as a one-dimensional construct and was measured with a single indicator. In the analysis at the level of a discussion thread, the indicator was the average of the ratings for all of the recommendations in the thread. In the analysis at the level of a recommendation, the ratings were used directly.

Inter-rater reliability for information integrity was found to be low, with the correlation just above 0.4 (see section 5.3); nonetheless, the correlations were statistically significant.

In all tests reported in Chapter 5 that included information integrity (at the level of a discussion thread or at the level of an individual recommendation), no statistically significant relationships between information integrity and other constructs were discovered. Moreover, a one-way ANOVA test (see section 5.8.5) suggested that there were no statistically significant differences between the levels of information integrity in different discussion threads. Thus, none of the tests supported the approach to conceptualizing and measuring information integrity attempted in this study.

These results highlight the difficulty of conceptualizing information integrity in health support group online discussions. In my view, low inter-rater reliability was because of the ambiguity of how recommendations are stated in the course of the discussions. The exact nature of the recommendation is not necessarily clear enough for the expert coders to reliably judge if the recommendation is ultimately based on evidence-based knowledge.

5.9.2 Support of High-Level Hypotheses

The purpose of this study was formulated in section 1.6; the specific research questions were formulated in the same section as two high-level hypotheses. This section discusses the implications for these hypotheses of the results of the model testing conducted in Chapter 5.

5.9.2.1 H1: Knowledge Construction Results in Higher Perceived Information Quality

The hypothesis H1 that knowledge construction results in higher perceived information quality was tested in a number of models, with constructs at different levels.

2x2 model. In the model in section 5.7, knowledge construction was conceptualized as a two-dimensional construct with the dimensions of explicitation and evaluation, with perceived information quality also conceptualized as a two dimensional construct, with the dimensions of relevance and usefulness. Explicitation affected both relevance and usefulness with medium effect sizes, but evaluation had no effect on them. The effect of explicitation on usefulness was slightly higher than the effect on relevance. The model had a substantial discriminant validity problem with relevance and usefulness highly correlated (as discussed in sections 5.6 and 5.9.1).

1x1 model. In the model in section 5.8.1, both knowledge construction and perceived information quality were conceptualized as one-dimensional constructs. Knowledge construction affected perceived information quality with a medium effect size. However, the explanatory power of the model (in terms of the amount of variance explained) was somewhat lower than the explanatory power of the model in section 5.7 (labelled as "2x2 model" in this section).

2x1 model. In the model in section 5.8.2, knowledge construction was conceptualized as a two-dimensional construct with the dimensions of explicitation and evaluation, with perceived information quality treated as a one-dimensional construct. Explicitation affected perceived information quality with a medium effect size, but evaluation had no effect on perceived information quality. The explanatory power of this model was similar to the model in section 5.7 (labelled as 2x2 model in this section) and higher than for the 1x1 model. On the other hand, the 2x1 model did not have the discriminant validity problems of the 2x2 model. Thus, of the models considered, the 2x1 model fitted the data the best.

Overall, all of the models analysed in Chapter 5 (including the three models discussed in this section) supported the H1 hypothesis. The models treating knowledge construction as a separate construct offered additional insight at the mechanism behind

the H1 hypothesis. The knowledge construction activities that contributed to perceived information quality were at a low level of knowledge construction (the activities represented by the explicitation construct). The results were consistent with a view that higher level knowledge construction activities (the activities represented by the evaluation construct) had no effect.

Thus, refinement (improving or perfecting by pruning or polishing), elaboration (developing in intricate and painstaking detail), clarification (interpretations that remove obstacles to understanding), confirmation (additional proofs that some facts or hypotheses are correct), and negation (statements that are a refusal or denial of some other statement) contributed to higher perceived information quality of the transcripts of health support group online discussions. Conversely, critical discussion (discussion characterized by careful evaluation and judgment), argumentation (a discussion in which reasons are advanced for and against some proposition or proposal), reasoning (presentation of reasons and arguments), and justification (defending or explaining or making excuses for by reasoning) did not. In other words, clarifying the facts contributed to perceived information quality, but reasoning about the facts did not. This is consistent with the findings by Kanuka and Anderson (2007) (discussed in section 2.6.2.1), who found little evidence of high levels of knowledge construction in online discussions by professionals and interpreted the findings by suggesting that high-level knowledge construction activities are not as relevant in online discussions in non-educational contexts as they are in educational contexts.

Clarification of facts contributes to perceived information quality in terms of making the facts more accessible, but reasoning and argumentation from a variety of perspectives do not necessarily clarify the facts and might even obscure them. Lasker et al. (2005) found that the main rationale for people to participate in health support group online discussions was gaining access to information (discussed in section 2.3). High levels of knowledge construction do not necessarily make it easier to access information, but clarification and refinement (lower level knowledge construction activities) do make it easier to access information.

5.9.2.2 H2: Knowledge Construction Activities in Health Support Group Online Discussions Result in Greater Prevalence of Evidence-Based Knowledge

Prevalence of evidence-based knowledge was conceptualized as information integrity. The hypothesis H2 that information integrity is affected by knowledge construction was tested in a number of models (see sections 5.7 and 5.8). The strongest test was reported in section 5.8.5, where the hypothesis that evaluation and explicitation (the dimensions of knowledge construction) affect information integrity was tested at the level of a recommendation. In all tests, the hypothesized effects on information quality were not found to be statistically significant. Thus, the results were consistent with the view that knowledge construction does not affect information integrity (and thus, knowledge construction does not result in greater prevalence of evidence-based knowledge).

A possible reason for this finding was that the operationalization of information integrity was not valid enough in view of low inter-rater reliability (see section 5.3 for a detailed discussion of this aspect).

Taking a view that the effects on information integrity were not discovered because there were no effects (rather than because of problems with the measure, or because the statistical power of the study was not sufficient), this suggests that knowledge construction activities in health support group online discussions are not robust enough to distinguish high-quality recommendations from recommendations that are doubtful or even risky.

5.10 Summary

This chapter presented the results of data analysis.

Descriptive statistics (minimum, maximum, mean, and standard deviation for each item, separately by coder) and inter-rater reliability (in terms of Pearson's correlation and Spearman's rho) were presented. Inter-rater reliability was the highest for the dimensions of knowledge construction, followed by the relevance and usefulness dimensions of perceived information quality. Inter-rater reliability for the understandability dimension of perceived information quality was so low (the

correlation between ratings by different coders was not statistically significant), that the construct was excluded from further analysis. Inter-rater reliability for information integrity was also low, but was judged to be high enough to retain it in further analysis. For the overwhelming majority of the indicators, the inter-rater reliability was below the threshold of 0.7; nonetheless, this was consistent with similar prior studies.

Normality tests suggested that the data were close enough to normal (according to kurtosis and skewness values); however, the Shapiro-Wilk's multivariate normality test suggested that the data were not multivariate normal.

Preliminary testing of convergent and discriminant validity using exploratory factor analysis suggested possible discriminant validity issues between the dimensions of knowledge construction (explicitation and evaluation) and between the dimensions of perceived information quality (relevance and usefulness).

The results of measurement model analysis were overall consistent with EFA results; nonetheless, the discriminant validity issues were minor enough to allow the analysis of the structural model.

Structural model analysis suggested that explicitation affected both relevance and usefulness; the rest of the hypotheses were not confirmed (no effects were discovered for evaluation, and no antecedents—for information integrity).

Post hoc analyses involving variations of the research model suggested that the best model fit (in terms of maintaining the explanatory power while avoiding discriminant validity issues) is achieved in a model with explicitation and evaluation treated as separate constructs, but with relevance and usefulness combined in a one-dimensional perceived information quality construct.

In a separate post hoc analysis, coders' conceptions of the target constructs were modelled explicitly, resulting in a hierarchical model. Coders' conceptions were found to be close enough to the underlying constructs to result in a good fit of the model.

The results of a post hoc analysis involving testing variations of the model at the level of a recommendation were consistent with the initial analysis in not discovering any

relationships. Moreover, a one-way ANOVA analysis suggested that there was no significant difference between the levels of information integrity in different threads.

Data analysis results across all of the analyses conducted were summarized and interpreted in section 5.9. The results suggested that knowledge construction in health support group online discussions improves perceived information quality, but does not affect information integrity. Thus of the two high-level hypotheses (H1 and H2) in the high-level research model in Figure 2, only the hypothesis H1 was confirmed.

Chapter 6. Discussion and Conclusion

6.1 Introduction

This chapter summarizes the results of the present study and discusses their implications. An overview of the main highlights of the study is followed by a discussion of the implications of the results for theory and for practice. The chapter concludes by stating the limitations of the research, making suggestions for future research, and by presenting the conclusions.

6.2 Overview of the Study

The problem addressed by the present study was the lack of objective evidence that health support group online discussions contribute to better quality of health-related information available online. More broadly, there is a dearth of studies of the effects of knowledge construction activities in online discussions in non-educational contexts.

The study was motivated by the evidence that patients are increasingly relying on online sources for health-related information and by the evidence that the majority of the participants in online discussions in non-educational contexts are lurkers and therefore access the content of the discussions merely as information available online.

The importance of pursuing the present study is highlighted by the ongoing transition to patient-centred health care—patients are expected to take a more active role in managing their conditions, and thus, need access to high-quality health-related information. Moreover, the research on health literacy has suggested that health-related information currently available online often does not match the levels of health literacy and the cultural preferences of its intended consumers. In health support online discussions, patients are directly involved, and therefore the resulting discussion transcripts are likely to fit the patients' levels of health literacy and cultural preferences. Nonetheless, health support group online discussions can positively contribute to the transition to patient-centred health care only if the information resulting from the

discussions is of high quality. Therefore, it is important to study how information of high quality emerges in the discussions.

A number of studies conducted in the context of education have suggested that knowledge construction is an important aspect of online discussions. The concept of knowledge construction captures the aspects of online discussions that are particularly relevant to learning. When patients are accessing information generated in health support online discussions, they are learning about health-related interventions and health behaviours. Therefore, knowledge construction is also highly relevant to health support group online discussions.

The present study focused on knowledge construction in health support group online discussions and on the quality of information of the resulting transcripts, with the quality of information conceptualized as perceived information quality (the quality of the information as perceived by the information consumers) and information integrity (the quality of the information from the perspective of the experts). The research questions were formulated as two high-level hypotheses (see the high level research model in Figure 2):

H1: Knowledge construction results in higher perceived information quality.

H2: Knowledge construction activities in health support group online discussions result in greater prevalence of evidence-based knowledge.

The prior research on knowledge construction and on information quality was reviewed to conceptualize and to operationalize knowledge construction, perceived information quality, and information integrity, thus enabling hypothesis testing.

The literature review in the present study uncovered numerous prior studies of knowledge construction in online discussions conducted in the context of education; the conceptualizations and the operationalisations of knowledge construction presented in these studies were synthesized, and the resulting synthesis formed the basis of the conceptualization and the operationalization of knowledge construction in the present study. The present study viewed knowledge construction as a two-dimensional concept with the dimensions of explicitation and evaluation. Explicitation corresponds to lower-level knowledge construction activities, such as refinement, elaboration, clarification,

confirmation, and negation of statements in the discussion. Evaluation corresponds to higher-level knowledge construction activities, such as critical discussion, argumentation, reasoning, and providing justifications. Explicitation and evaluation were operationalized by analysing their definitions in the literature and by focusing on the most important keywords.

Perceived information quality was conceptualized and operationalized based on the study by McKinney et al. (2002). Of the dimensions of perceived information quality suggested by McKinney et al., the dimensions of relevance, understandability, and usefulness were judged to be particularly relevant to the present study.

Information integrity was conceptualized as the extent to which the recommendations for health-related behaviours and interventions suggested by the discussions are recognized by experts as rooted in evidence-based medicine. Thus, information integrity was conceptualized as a one dimensional construct.

The conceptualizations of knowledge construction, perceived information quality, and information integrity formed the basis of the detailed research model of the present study, in which the high-level hypotheses H1 and H2 were elaborated at the level of the individual dimensions of the high-level constructs (see Figure 8). The research model comprised eight hypotheses. Both evaluation and explicitation were hypothesized to affect the dimensions of perceived information quality, relevance, understandability, and usefulness; moreover, both evaluation and explicitation were hypothesized to affect information integrity.

The present study adopted a positivist research paradigm and a quantitative approach to data collection and analysis. Positivism and quantitative research were better suited to address research questions of the study (formulated as hypotheses) than interpretivism or qualitative research. Moreover, I believe that a study demonstrating the effects of knowledge construction relying on a paradigm that claims objectivity may contribute to providing basis for further research relying on a range of paradigms and approaches, such as, for example, qualitative research exploring the context-bound nuances and the subjective meanings of knowledge construction and its effects.

Content analysis was adopted as a quantitative research technique that maximizes ecological validity—content analysis allows to study real-world discussions without interfering with the discussions. Moreover, content analysis was feasible to conduct within the resource constraints of the study. The approach to content analysis by Krippendorff (2004) was followed.

To test the research model, the content of the discussion in a health support group devoted to weight management was used (the content was available in the public domain). The unit of analysis was a discussion thread. The content was sampled for the threads initiated between the first week of August 2010 and the third week of October 2010 (1,390 discussion threads overall).

The choice of a discussion thread (rather than a message or a sentence) as a unit of analysis distinguished the present study from most of the prior studies of knowledge construction relying on content analysis. I chose a discussion thread as a unit of analysis because the pattern of knowledge construction activities such as elaboration, negation, argumentation, or reasoning, develops within a discussion thread, rather than within a message (and even less so, within a sentence). Often, messages in online discussion make sense only within the context of the discussion thread. Messages within the same discussion thread share the same mood and the same approach to discussion, as it develops over the discussion. Therefore, it makes more sense to associate the level of knowledge construction with the whole discussion thread rather than with a single message. The demonstration of the viability of using a discussion thread as a unit of analysis in studying knowledge construction in online discussions can be seen as a contribution of the present study to methodology.

Following the approach to content analysis by Krippendorff (2004), suitably qualified coders were employed as research assistants and their characteristics were documented to enable replications of the present study. Coding schemes were developed based on the operationalisations of the research model constructs (Likert scales were used in the coding schemes, to make is possible to use structured equation modelling for hypotheses testing). To enable replications, coders coded the content based on the coding schemes, and no undocumented coder training was conducted. To minimize the risk of common method variance, different coders were used to code constructs

hypothesized to be related. Moreover, to allow the assessment of inter-rater reliability, each construct was coded independently by two coders.

Discussion threads were coded for knowledge construction and perceived information quality directly, in a single step. Information integrity, however, was coded in two steps. First, coders with university education but with no expertise in health extracted health-related recommendations from the discussions. Second, health practitioners with nursing degrees employed as expert coders assessed the extracted and summarized recommendations for adherence to evidence-based medicine. The two-step procedure was based on the Delphi technique and was needed to minimize the time commitment of the expert coders (their time is scarce and expensive). The codes for individual recommendations were aggregated to obtain information integrity ratings at the level of a thread. The procedure for operationalizing information integrity was developed in the present study and can be seen as a contribution of the present study to methodology.

Of the discussion threads sampled, only typical threads were used for hypotheses testing. The thresholds determining whether a thread was treated as typical were based on examining thread distributions in the sample and to an extent were based on my subjective judgment (threads with three to five contributors and containing from five to seven messages were considered typical). Typical threads tended to be large enough to allow knowledge construction but small enough to be codable as single units.

Of the 1,390 threads sampled, 501 were typical. The study relied on PLS SEM for hypothesis testing, and the methodological literature and the prior studies relying on PLS suggested that approximately one hundred data points are needed to test the research model of the present study. Therefore, 120 of the typical threads were selected at random for coding. Only 108 of these threads contained health-related recommendations, and thus could be used to test the research model. These threads were coded for all of the constructs of the research model and the resulting codes were used to test the model.

The results for inter-rater reliability were mixed. For all of the indicators of explicitation and evaluation (the dimensions of knowledge construction), the correlation between the codes obtained by different coders was either above or very close to the threshold of 0.7 suggested by Nunnally and Bernstein (1994); moreover,

the correlation values corresponded to large effects in terms of the Cohen's (1992) classification of effect sizes, and the correlations were statistically significant at $p<0.001$. As to the relevance and usefulness dimensions of perceived information quality, even though the correlation values were below the threshold value of 0.7, they corresponded to large effect sizes according to Cohen and were statistically significant at $p<0.001$. For the understandability dimension, however, the inter-rater reliability was very poor—for most of the indicators the correlations were not even statistically significant. For information integrity, the inter-rater reliability was similar to relevance and usefulness. Based on the results for inter-rater reliability, the understandability dimension was removed from the model (resources required to iteratively develop the coding schema for understandability to improve its reliability were not available). For the remaining constructs, the levels of inter-rater reliability obtained in the present study were similar to the prior studies of knowledge construction in online discussions that employed independent coders. In model testing, the codes by the independent coders for each of the indicators were added up, resulting in a single value for each indicator.

The study employed PLS SEM to test the measurement and the structural model. In addition, EFA analysis was conducted as a preliminary test of the measurement model.

In an EFA analysis with the number of factors set to four, all of the indicators of explicitation and evaluation, the dimensions of knowledge construction, loaded on a different factors (with the exception of the negation indicator of explicitation, which loaded the highest on the same factor as the indicators of evaluation). In contrast, all of the indicators of relevance and usefulness, the dimensions of perceived information quality, loaded on the same factor. Nonetheless, in an EFA analysis involving the indicators of perceived information quality only the indicators of relevance and usefulness loaded on different factors. Overall, EFA analysis supported the treatment of knowledge construction as a two-dimensional construct, but the result was ambiguous for validity between relevance and usefulness. (Information integrity had only one indicator and therefore was not included in the EFA analysis.)

In PLS measurement model analysis, the discriminant and the convergent validity were confirmed for all of the constructs. All indicators loaded on their own constructs the

highest and higher than the threshold value of 0.7 (or very close to the threshold). The Cronbach's alpha and the composite reliability values were above the threshold of 0.7, further suggesting the convergent validity of the measures, and square roots of AVE for each construct were higher than the correlations of the construct with other constructs in the model. The correlation between relevance and usefulness, however, was very higher, putting the discriminant validity between relevance and usefulness somewhat in doubt (even though the rest of the criteria for discriminant validity were met). Thus, even though the PLS measurement model analysis provided stronger support for discriminant validity between relevance and usefulness than the EFA analysis, the support was ambiguous in both of the analyses.

The results of the PLS SEM analysis of the structural model suggested that explicitation (the dimension of knowledge construction corresponding to lower-level knowledge construction activities) affects both of the dimensions of perceived information quality, with medium effect sizes according to the criteria suggested by Kline (1998). Evaluation (corresponding to higher-level knowledge construction activities) did not affect either perceived information quality or information integrity. Moreover, information integrity was not affected by either of the dimensions of knowledge construction. Thus, the results of the structural model analysis suggested that knowledge construction contributes to perceived information quality (supporting the H1 high-level hypothesis of the present study). However, the results of the structural model analysis provided no support for the high-level hypothesis H2 (that knowledge construction contributes to information integrity); the results were consistent with a view that knowledge construction does not affect information integrity.

To further explore the meaning of the data, a series of post hoc analyses were conducted. PLS SEM analyses of variations of the research model in which knowledge construction and/or perceived information quality are treated as one dimensional suggested that the model with the separate dimensions of knowledge construction (explicitation and evaluation) and with a single dimension of perceived information quality (combining relevance and usefulness) fits the data the best: It had no discriminant validity issues and, at the same time, retained the explanatory power of the initial research model (the model with both knowledge construction and perceived information quality treated as multidimensional).

The analysis of a model with an added hypothesis that explicitation affects evaluation (thus, lower level knowledge construction activities enabling higher level knowledge construction activities) provided a strong support for the added hypothesis (with the effect size strong according to the criteria suggested by Kline, 1998). This result further supported the validity of the overall analysis.

Post hoc analyses at the level of a recommendation (with a larger number of cases that the analyses at the level of a thread, and, therefore, with higher statistical power) still did not detect any effects on information integrity. Moreover, analyses of the data separately for each coder that coded information integrity (to explore the possibility that one of the coders coded incorrectly) were also conducted and no effects on information integrity were discovered. The results of all of the analyses were consistent with a view that knowledge construction does not affect information integrity (even though the possibility that the lack of effects detected was because of problems with the coding schema cannot be excluded).

Overall, the results of all of the analyses consistently supported the high-level hypothesis H1 (that knowledge construction contributes to perceived information quality) and did not support the high-level hypothesis H2 (that knowledge construction contributes to information integrity). Moreover, lower-level knowledge construction activities (explicitation) mattered more than higher-level knowledge construction activities (evaluation).

Thus, the results suggest that knowledge construction in health support group online discussions contributes to better quality of health-related information available online in terms of perceived information quality. Lower-level knowledge construction activities—refinement, elaboration, clarification, confirmation, and negation—result in information that appears to be more relevant and more useful. Nonetheless, higher-level knowledge construction activities—critical discussion, argumentation, reasoning, and justification—do not contribute to perceived information quality.

Finally, knowledge construction activities (either high- or low-level) do not contribute to information integrity—the extent to which the information is valid in terms of evidence-based medicine. It is important to emphasize that this finding does not suggest that the information available via health support group online discussions lacks

integrity, but only that the knowledge construction activities per se do not result in higher (or lower) information integrity.

6.3 Implications of the Study

6.3.1 Implications for Theory

This section describes the contributions of the present study to theory and to the body of literature.

6.3.1.1 Extending the Nomological Framework Around Knowledge Construction

By demonstrating the effect of knowledge construction on perceived information quality, this study contributed to building a nomological framework around the knowledge construction construct. The nomological framework around the knowledge construction construct simplifies the formulization of valid measures of knowledge construction, as new measures can be tested by using models from the framework.

In contrast to the present study's focus on the consequences of knowledge construction, prior quantitative studies that involved testing hypotheses regarding relationships between knowledge construction and other constructs focused on the antecedents of knowledge construction (possibly, because all of these studies were conducted in education context, where high levels of knowledge construction are seen as a desirable outcome in themselves). The present study extended the nomological framework by considering the outcomes of knowledge construction in online discussions. In the rest of this section, I synthesize what is known about antecedents and outcomes of knowledge construction by formulating an integrated model based on the results currently available, including the results of the present study.

Involvement in online discussions (in terms of intensity and in terms of experience) has been found to result in higher levels of knowledge construction. Schellens and Valcke (2005), in a study of educational science freshmen' contributions to a discussion group, tested the hypothesis that higher level of involvement results in higher level of

knowledge construction. The hypothesis was confirmed. In the same study, they tested the hypothesis that the individuals taking part in an online discussion improve their knowledge construction ability as the discussion progresses, so that messages posted later in the discussion correspond to higher level of knowledge construction, which was not confirmed; however, in a similar study of educational science freshmen (Schellens and Valcke, 2006), this hypothesis was confirmed.

Moderation and scaffolding of online discussions have been found to result in higher levels of knowledge construction. De Wever et al. (2009; 2010), based on content analysis of online discussions by first-year students in instructional sciences, found that assigning roles to discussion contributors (roles such as starter, moderator, and source searcher) promoted knowledge construction. In a recent study, Rienties et al. (2012) compared knowledge construction in economics students' discussions using a system that offered scaffolding of the learning process by providing explicit reminders and guidance to facilitate productive planning with knowledge construction in discussions using a system that did not provide such scaffolding. Rienties et al. found that scaffolding promoted knowledge construction.

In view of the results of the present study, as summarized in the diagram in Figure 30, it should be expected that high levels of involvement in online discussions and involvement over time (which can be seen as contributor behaviours), along with scaffolding and moderation (which can be seen as support activities), result in information of higher perceived information quality (in terms of perceived usefulness and relevance) captured in the discussion transcript. The diagram highlights the contribution of the present study—in contexts other than in education, knowledge construction activity is not an aim in itself. The present study demonstrated that knowledge construction results in a benefit that is relevant to practice, namely, information of better quality. By combining the result of the present study with the results of the prior studies, one can see how the quality of information generated via knowledge construction in online discussions can be improved: scaffolding and moderation can be provides by suggesting to the contributors specific dialogue steps or by assigning contributor roles. Moreover, one can encourage contributor involvement intensity (e.g., in terms of the numbers of good-quality contribution) and prolonged involvement by providing rewards, such as assigning visible labels describing

contributor status (such as “beginner” to less active contributors and “expert” to contributors active over a period of time).

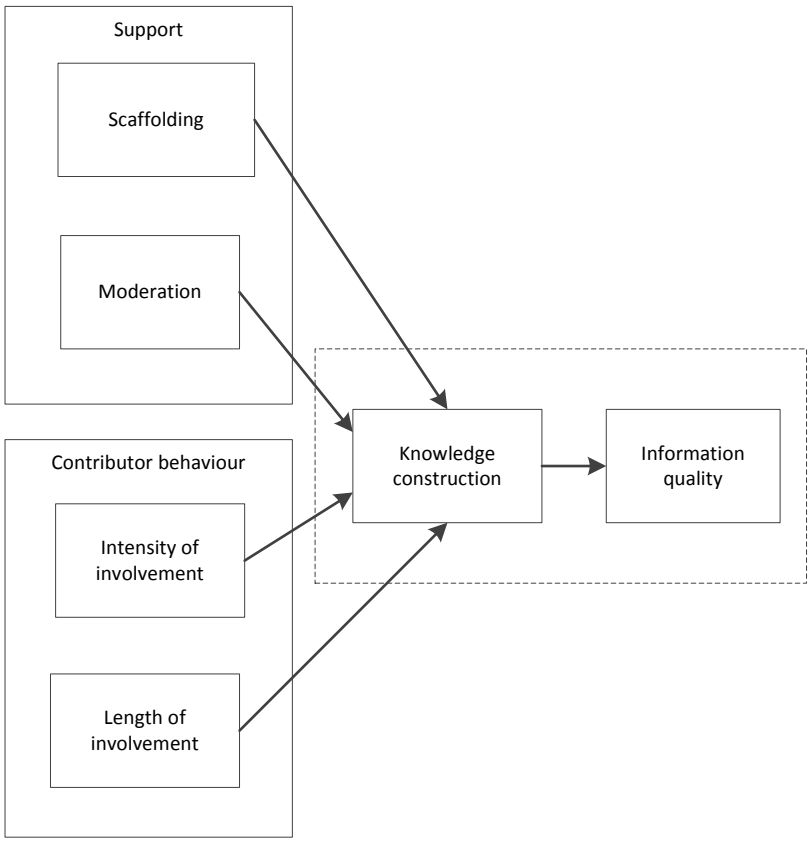

Figure 30. Nomological framework around knowledge construction in online discussions. The part of the framework addressed by the present study is shown in a box with a dashed border.

6.3.1.2 The First to Conceptualize and to Measure Knowledge Construction as a Multidimensional Variable

In prior content analysis studies of knowledge construction (reviewed in section 2.5), knowledge construction has been measured as a categorical variable describing the type of a knowledge construction activity or as an ordinal one-dimensional variable describing the level of knowledge construction (for example, measured as low, medium, or high in the study by Järvelä and Häkkinen, 2002). The present study was the first to conceptualize and to measure knowledge construction as a multidimensional variable with the dimensions reflecting the extent of activities at different levels of knowledge construction. The study demonstrated the advantages of using such a

conceptualization and measure in terms of understanding the effects of knowledge construction. A study relying on a one-dimensional measure of knowledge construction would not be able to detect that even though knowledge construction affects perceived information quality, the effect is coming from explicitation (low-level knowledge construction activities) rather than from evaluation (high-level knowledge construction activities). One dimensional measures reflecting the presence of more sophisticated knowledge construction activities may be appropriate in education, where sophistication of knowledge construction discourse may be seen as an end of itself. In other domains, such as those involving specific pragmatic goals, multidimensional measures are more appropriate.

The finding that knowledge construction is a multidimensional construct suggests the need for a "middle ground" in knowledge construction research. In past research, knowledge construction has been studied primarily by qualitative methods, resulting in rich descriptions of knowledge construction activities and outcome. A recent example of qualitative approach is the study by Secko, Tlalka, Dunlop, Kingdon, and Amend (2011), which employed qualitative content analysis of commentary to online content of the Canadian Globe & Mail newspaper to describe how readers develop health and science related topics initiated by the newspaper articles via online discussions. Secko et al. concluded that such online discussions result in a shift of power from newspaper journalists to the readers. Qualitative studies present knowledge construction as a multifaceted, rich phenomenon. At the same time, the prior quantitative studies of knowledge construction antecedents and outcomes (as reviewed in section 2.5 and the preceding section, section 6.3.1.2) have regarded knowledge construction as a one-dimensional construct (from practice perspective, adopting the position that the more knowledge construction, the better). This can be contrasted with the quantitative studies of information quality, which introduced multiple dimensions of information quality. For example, McKinney et al. (2002) considered the relevance, understandability, reliability, adequacy, scope, and usefulness dimensions of information quality, and the earlier review of literature on data quality by Wang and Strong (1996) listed believability, accuracy, objectivity, and reputation as intrinsic data quality dimensions; value-added, relevancy, timeliness, completeness, and appropriate amount of data as contextual data quality dimensions; interpretability, ease of understanding,

representational consistency, and concise representation as representational data quality dimension; and accessibility and access security as accessibility data quality dimensions. Clearly, these dimensions represent data quality from a broad variety of perspectives.

The results of the present study suggest that similarly to information quality, knowledge construction in quantitative studies should be described by multiple dimensions, presenting a rich view of the knowledge construction process, thus bridging the gap between qualitative and quantitative approaches to studying knowledge construction.

Conceptualizing knowledge construction as a multidimensional construct allows one to understand better the role of knowledge construction in specific contexts, such as the context of health support group online discussions. In different contexts, different dimensions may be of relevance.

The drawback of conceptualizing knowledge construction as a multidimensional construct in quantitative research is that this would result in research models that are not parsimonious, and thus vulnerable to capitalization of chance in model testing (capitalization of chance refers to a relationship found to be statistically significant because of the particular features of the data set, rather than because the theory suggesting the relationship is valid, see section 5.7). To guard against capitalization of chance, studies of antecedents and consequences of knowledge could use models that include only the dimensions of knowledge construction that are expected to be of importance under the particular circumstances in which he studies are conducted (thus, using a contingency approach to including the dimensions).

6.3.1.3 Confirmed the Relevance of Knowledge Construction Beyond the Domain of E-Learning

By conducting a study of knowledge construction in a non-educational domain, this study confirmed the relevance of knowledge construction beyond the domain of e-learning at educational institutions, and contributed to paving a way for further studies of knowledge construction in non-educational domains, such as knowledge management.

The results of the present study suggest that the concept of knowledge construction is relevant in the study of the uses of collaborative software in non-educational contexts. Over the recent years, there has been a sustained interest in the use of software systems enabling spontaneous online collaboration (and thus, potentially enabling online knowledge construction), such as discussion forums, blogs, wiki, and social media, in organizational contexts. Mehta (2008) conducted a multiple case study of knowledge management at global software companies that highlighted the role of discussion forum software as an enabler of knowledge management. Paroutis and Al Saleh (2009) conducted a series of interviews of bloggers at a major (undisclosed) technology and services corporation to investigate the enablers and the barriers to the use of online collaboration tools, such as blogs and discussion forums, for knowledge management; they highlighted the importance of management support and trust. Magnier-Watanabe, Yoshida, and Watanabe (2010) conducted a survey of Japanese workers to investigate the effects of intranet-based social networking services on productivity and found that social networking has mild positive effect on access to knowledge. Vuori and Okkonen (2012) conducted a survey at two industrial companies to investigate the motivation behind use of social media for knowledge sharing, and found that the desires to promote organizational goals and to help colleagues were the main drivers. Chua and Banerjee (2013) conducted qualitative content analysis to understand the efforts of the Starbucks corporation to involve its customers in knowledge sharing via social media to transform the customers into active contributors of innovation and concluded that such customer involvement could be a game changer for traditional (not online) businesses. Yet, unlike in the field of education, none of the studies of collaborative software use at organizations focused on knowledge construction (with the exception of the relatively old study by Kanuka and Anderson , 2007, introduced in section 2.6.2.1), even though the broader aspect of organizational knowledge creation (subsuming knowledge construction via online discussions) is a central issue in knowledge management (Dalkir, 2011).

The finding of the present study that knowledge construction in a non-educational context results in better perceived information relevance and usefulness of the discussion transcripts highlights the relevance of knowledge construction in non-educational domains. In particular, the finding that only explicitation, but not

evaluation, had effect may be attributable to the non-educational context of the study; in educational contexts the contributors are likely to have higher levels of literacy and are explicitly encouraged and taught how to engage in evaluation (because critical thinking is viewed as a desirable outcome of learning). Therefore, the evaluation knowledge construction activities may be more productive in educational contexts, but explicitation knowledge construction activities are may be more productive in non-educatonal contexts.

Therefore, the findings of the present study not only confirm the relevance of studying knowledge construction in online discussions conducted in non-educational contexts, but also emphasize the need for such studies—findings for the effects of knowledge construction in education cannot be mechanically generalized to other domains. In particular, in non-educational contexts, attaining high levels of knowledge construction is not an aim in itself. Moreover, the contributors in non-educational contexts may have lower levels of general literacy and literacy relating to the topics under discussion, such as health literacy relating to health support group online discussions, making higher-level knowledge construction activities less productive.

6.3.2 Implications for Practice

This section describes the implications of the present study to practice.

6.3.2.1 Online Discussions Do Have Value in Terms of Enabling Individuals to Access Health-Related Information

Both active contributors to asynchronous health-related online discussions and, particularly, lurkers are exposed to transcripts of the discussions. The finding of this study that explicitation results in better perceived information quality suggests that such online discussions do have value in terms of enabling individuals to access information. Thus, this research contributes to the research stream of the studies of the value of health support group online discussions reviewed in section 2.3 of this thesis.

The results of the present study are supported and illustrated by the studies by Mankoff, Kuksenok, Kiesler, Rode, and Waldman (2011) and Fraidaki, Pramatari, and Doukidis (2012). Mankoff et al. conducted a study of the use of online resources by people with a

chronic illness (Lyme disease), which involved a survey and a series of in-depth interviews. Mankoff et al. uncovered a pattern of patients coming to terms with conflicting viewpoints expressed on the Internet, evolving identity and beliefs about their illness, with some of the patients led to create their own content and advising others. In terms of the conceptualization of the present study, this pattern can be interpreted as knowledge construction. Fraidaki et al., based on a survey of Internet users, found that 63% of the users obtain health-related information from consumer review pages, and that such information influences their health-related decisions. Review pages are structured similar to discussion forum discussions, with contributors engaging in online dialogue about the product reviewed. Therefore, the Fraidaki et al. findings are consistent with the outcome of the present study—discussions via review pages result in better quality information, and such information influences the lurkers (users surveyed in the study by Fraidaki et al. did not post reviews themselves, but accessed posts by others).

The results of the present study help to interpret the results by DeAndrea and Anthony (2013). DeAndrea and Anthony, based on a survey of non-institutionalized US adult population, found that large numbers of depression sufferers in the US seek support via online chat and discussions. DeAndrea and Anthony found the results of their study inconclusive in terms of whether or not such online activity is of benefit to the patients in allowing them to better manage their conditions. The results of the present study are congruent with a view that online knowledge construction by depression sufferers is likely to result in better quality information (in terms of perceived relevance and perceived usefulness), and thus be of benefit to the discussion participants and for lurkers.

The importance of the results of the present study is highlighted by the findings by Waters, Canfield, Foster, and Hardy (2011). Waters et al. studied the engagement of students with a university health centres' social networking sites by content analysing online material in terms of the Kent and Taylor's dialogic principles of communication (Kent & Taylor, 1998). They found that when a health centre social networking site attracts a large number of followers online, it results in high levels of activity and complexity of online dialogues by the students. In terms of the theoretical framework employed in the present study, such dialogues can be seen as knowledge construction.

The results by Waters et al. suggest that as health organizations become more active in engaging patients and populations at risk online, they initiate intense knowledge construction activity. The results of the present study suggest that such activity enhances the effects of the health organizations' online presence, with knowledge construction leading to health-related information with higher perceived relevance and perceived usefulness.

6.3.2.2 Moderators Should Promote Clarifications and Refinements Rather Than Arguments

Moderators are discussion contributors charged with supporting and directing the flow of the discussions (Salmon, 2003). Moderators need not be subject member experts.

The results of this study suggest that to make the health support group online discussions more effective, moderators should encourage clarifications and refinements (and, overall, promote explicitation), rather than encourage arguments and justifications (and thus promote evaluation).

In his book, Salmon (2003), based on his ten years of experience in online teaching, proposed a five stage model of e-moderation involving the stages of (1) access and motivation, (2) online socialization, (3) information exchange, (4) knowledge construction facilitation, and (5) meta-learning and reflection. Wright (2009), based on two case studies of government-run discussion sites, identified the following e-moderator roles: "greeter", "conversation stimulator", "conflict resolver", "summarizer of debates", "problem solver", "supporter", "welcome", "cybrarian", "open censor", "covert censor", and "cleaner" (p. 236). It is notable that Wright put more emphasis than Salmon on restrictive actions by the moderator to protect the discussion from misuse (probably, because the case studies on which the Wright's moderator roles are based included the use of discussion forums as a media for expressing political views, rather than for constructive collaboration and mutual support).

The results of the present study are more in line with the model by Salmon (2003), but suggest a particular interpretation of the knowledge construction facilitation step. Namely, the moderator should promote the type of knowledge construction that matches the level of literacy of the discussion contributors—rather than promoting

evaluation, the moderator may need to focus on explicitation. In particular, at the health support group online discussion devoted to weight management studied in the present study, the most relevant dimension of knowledge construction was explicitation, a finding that may be generalizable to similar health support group online discussions. It is likely that explicitation matched the level of health literacy of discussion participants, making explicitation a productive knowledge construction activity (it is likely that the contributors were health literate enough to clarify, but not health literate enough to evaluate and critique effectively).

The suggestion that moderation and scaffolding should be employed at health support group online discussions is supported by the recent findings by De Wever et al. (2009, 2010) and Rienties et al. (2012) (these studies were already introduced in section 6.3.1.1). De Wever et al. found that assigning roles to discussion contributors results in higher levels of knowledge construction, and Rienties et al. found that the levels of knowledge construction in discussions using a system that provided scaffolding of the learning process were higher than the levels of knowledge construction using a system where such scaffolding was not provided. As already noted in section 6.3.1.1, when combined with the results of the present study, the results by De Wever et al. (2009, 2010) and Rienties et al. (2012) imply that by promoting knowledge construction, moderation and scaffolding ultimately result in information of higher perceived information quality generated in the discussion.

In terms of the practicalities of employing moderators for health support group online discussion, there is no need to employ registered health practitioners as moderators, because the role of moderators is to facilitate the discussions, rather than to contribute health-related knowledge or to evaluate knowledge contributions. Rather, regular contributors with sufficient experience could assume moderator roles, with the health support group online discussion functioning as a self-managing online community (for a discussion of self-managing online communities, refer to Preece & Maloney-Krichmar, 2003). The possibility and the practicality of involving health practitioners to ensure information integrity (rather than to facilitate the discussions) is discussed in section 6.3.2.3).

6.3.2.3 Knowledge Construction in Health Support Group Online Discussions Does Not Contribute to Information Integrity

The results of this study are consistent with a view that knowledge construction in health support group online discussions does not contribute to information integrity. Thus, it is desirable that qualified health practitioners take part in such discussions to promote health-related behaviours based on evidence-based knowledge, and to expose recommendations that have uncertain or even dangerous effects. It is unlikely that the discussion participants are going to evaluate the quality of the recommendations sufficiently by themselves. Indeed, if the recommendation provided in health support group online discussions have dangerous effects when implemented, the positive effects of lower-level knowledge construction activities on perceived information quality would only make these recommendations even more dangerous.

The finding of the present study that knowledge construction in health support group online discussions does not result in improved information integrity raises an alarm regarding the possible negative consequences of contributors and lurkers being exposed to health-related information that is not evidence based, and, therefore, possibly ineffective or even harmful. Similar concerns have been raised in multiple prior studies. Goldberg, Berman, and Gusberg (2011) analysed the content of information on abdominal aortic aneurysm available online, and found important omissions in such information. Similarly, Haigh and Costa (2013) conducted content analysis of information available online on assisted dying and found that only 14% of the web sites were highly accurate. Post and Mainous (2010) analysed the content of nutrition information for type 2 diabetes available online, and found that such information was often incomplete and not up-to-date. Jena and Goldman (2011), based on an analysis of secondary data from the US, found that higher Internet use is positively associated with prescription drugs abuse, thus suggesting that the information obtained online results in harmful behaviour. Powell, Inglis, Ronnie, and Large (2011) conducted an online survey to study benefits and challenges faced by the users of the UK National Health Service Direct web site and found that determining what to trust was a "common sense" activity, and thus, the users of the web site had to rely on their own judgment; thus, information on which online information to trust was not available online, leaving the users vulnerable. The results of the present study suggest that knowledge construction

activities at health support group online discussions do not result in critical assessments that would solve the problem highlighted by the Powell et al.'s study.

At the beginning of this section I suggested that online collaboration between patients and health practitioners is likely to result in high integrity health-related information. The outcome of the study by Rajagolan et al. (2011) lends support to this suggestion. Rajagopalan et al. conducted content analysis to compare the content of articles with patient-oriented cancer information provided by Wikipedia and by a professionally maintained database (Physician Data Query). Wikipedia articles were found to have accuracy and depth similar to the professionally edited database. Wikipedia articles are maintained by volunteers, and it would be reasonable to assume that both patients and health practitioners contributed to the Wikipedia articles; the Physician Data Query is maintained by experts. Thus, when laymen and health practitioners collaborate online, the resulting knowledge construction process may have a different dynamics comparing to when just laymen are present. In particular, the health practitioners' criticisms and evaluations may be highly effective, leading to high integrity information.

The result by Lee, Gray, and Lewis (2010) further emphasizes the need to manage information integrity in health support group online discussions, to ensure that medical decision making by contributors and lurkers is conducted on a solid basis. Lee et al., based on a survey of patients registered in the Pennsylvania Cancer Registry, concluded that Internet use by cancer patients let them to be more active participants in medical decision making. As follows from the Lee et al.'s conclusion, involvement in health support group online discussions is likely to make the contributors more active in making health-related decisions, magnifying the effect of any information lacking integrity that they may have encountered. This view is further supported by the results by Lee and Sundar (2012), who conducted a simulation study of the credibility of tweets with health information. They found that even though tweets by health practitioners were perceived to be more credible than tweets by laypersons, retweets by laypersons were perceived as more credible than tweets by health practitioners. This result emphasizes the influence of health-related information obtained from online interactions, as under certain circumstances contributions by laypersons may be more influential than contributions by health practitioners. Thus, information with no integrity generated in online discussions may have harmful consequences.

The results of the present study suggest that contributors and lurkers in health support group online discussions should be proactively made aware that not all health-related information found on the Internet is reliable (because interactions between laymen contributors do not result in information lacking integrity being exposed as such). The results of the study by Gauld (2011) suggest that there is a demand for such a service. Gauld conducted a survey of Australia and New Zealand residents to investigate their use of health-related information available online. Highlighting the importance of the Internet as a source of health-related information, 12.6% of Australian respondents and 19.5% of New Zealand respondents had used the Internet to obtain health-related information. Out of the respondents who had used the Internet to obtain health-related information, 90% believed Internet health information to be reliable, and 35% routinely checked the credentials of information suppliers. The relatively high percentage of users that checked the credentials of information suppliers in the study by Gauld suggests that there is a demand for the participation of health practitioners promoting evidence-based knowledge.

Direct participation of health practitioners in health support group online discussions may be problematic for the following reasons. First, health practitioners have a duty of care—they are required to proactively help people under threat (General Medical Council, 2013; Mansfield et al., 2011). When participating in health support online discussions, health practitioners may identify individuals under threat. Then, they would be under obligation to engage individuals to ensure that they get medical care. Thus, the practitioners' participation is likely to go beyond exchange of ideas and discussion, but is likely to put the practitioners into complex, contradictory, and unpredictable social situations, with the practitioners expected to act, but possibly unable to act because of the lack of physical contact with the individuals under threat and because the infrastructure and the resources available at hospitals and other designated places of care are not available at on-line discussions. It is likely that the practitioners would resist putting themselves in such a situation. (In fact, the simple perceived presence of the practitioners in the discussion may put the practitioners at risk, because some of the contributors or lurkers might assume that all information posted is approved by the practitioners, and raise complaints when such information is found to lack integrity.)

Second, there is a shortage of qualified health practitioners in New Zealand and throughout the world (Gorman, Horsburgh, & Abbott, 2009) because of factors such as aging populations in developed countries. Therefore, health practitioners may simply not have time to engage in health support group online discussions. Providing care to patients is likely to be higher in priority.

Thus, it may be difficult or impossible to directly involve health practitioners in health support group online discussions.

A possible solution could be to train some of the contributors to improve their health literacy to make them expert in particular aspects of healthcare, and to afford them a special status, visible to other discussion participants (along with appropriate disclaimers). Such contributors could get rewards for continued involvement in the discussions as trained experts (in fact, ongoing training to maintain their health literacy and knowledge could be a reward by itself). The obvious danger of such as solution is that discussion participants may perceive the participants with the status of trained experts as if they were health practitioners and rely too much on their advice, resulting in possible harm and even in legal problems, as activities of trained experts could be interpreted as malpractice.

Another solution would be to involve health practitioners in health support group online discussions indirectly. This could be done by using an approach similar to the one described by Nordfeldt, Angane-Lindberg, and Bertero (2012), who presented a case study of the use of an Internet portal devoted to diabetes management by health practitioners and by patients. Some of the practitioners never used the portal, but others used it frequently and promoted it to their patients. The portal contained documents contributed by the practitioners; and the practitioners felt safe to recommend the documents to their patients because the documents were written by suitably qualified experts. Some of the practitioners read patients' exchanges at the discussion forum, and found them a valuable source of information on the patients' needs.

The results of the present study suggest that in a situation similar to the one described by Nordfeldt, Angane-Lindberg, and Bertero (2012) (when both health practitioners and patients have access to a web portal including a discussion forum), health practitioners

should participate in knowledge construction by the patients interacting at the discussion forum. This could be achieved by making available via the information portal new documents, reflecting the needs and the level of health literacy expressed by the patients at the discussion forum. Patients, in their turn, could discuss the content of the documents provided by the health practitioners, thus continuing the knowledge construction.

By posting documents at the portal, rather than participating in the discussions directly, health practitioners would clearly indicate the information that is evidence-based (in terms of the research model of the present study, information with integrity). At the same time, health practitioners would shield themselves from any claims that information available at the discussion forum leads to harm, because the practitioners do not contribute information to the discussion forum. (This approach raises an ethical dilemma—if something posted at the discussion forum may lead to harm, should the practitioners take action to remove it? Such action, though, could make them vulnerable to complaints, as the only "safe" policy appears to not interfere with the discussion forum at all.)

Moreover, to further reduce the demands on health practitioners' time, the practitioners could receive summaries of discussions compiled over a period of time, rather than read discussion forum posts directly (in fact, the procedure used to assess information integrity in the present study, see section 4.8.6, can be seen as an early prototype of such a process). Such summaries would not include sufficient details and would not be current enough to put the practitioners in situations where they need to exercise the duty of care with respect to specific individuals. Such summaries, however, should contain sufficient information to inform the participants about the levels of health literacy of health support group online discussion contributors and about the medical issues important to them, thus allowing the practitioners to take part in the knowledge construction process by making suitable documents available via the portal. Of course, information about such documents becoming available could be fed into the discussions (possibly, as a regular update about the state of the portal, generated automatically by the system). Such information is likely to result in further discussions, thus closing the loop. In terms of the distinction between explicitation and evaluation, the documents would strengthen the evaluation process in the discussion. The results of the present

study suggest that evaluation activities by health support online group contributors themselves are not effective; therefore, the strengthening of the evaluation process is highly desirable.

The suggestions (made in the present study) for designing a health information portal allowing health practitioners to get involved in knowledge construction at health support group online discussions should be considered in context of the plans to set up personal health portals in the New Zealand Health IT Board national health IT plan (2010). The health information portal proposed in the present study could be implemented within the same framework.

6.3.2.4 Implications for Education

Even though the focus of the study was on health support group online discussions, and not on discussions in education, the results of the study have implications for the use of discussion forums in health education and in education in general. Student involvement as discussion forum contributors may be low. As Internet users in general, many of the students act as lurkers and limit their involvement to reading postings by others (Kucuk, 2010; Nagel, Blignaut, & Cronje, 2009; Sloep & Kester, 2009). The model in Figure 30 suggests that knowledge construction in discussion forum discussions will result in relevant and useful information. In particular, because lurkers are similar to contributors, issues discussed and resolved by the contributors are likely to also be relevant to the lurkers, and solutions useful to contributors are also likely to be useful to lurkers. Course materials, however, are not necessarily focused on issues that are immediately relevant to the students, and may be using language that is difficult for the students (e.g., assuming prior knowledge that the students do not have). Thus, discussion forum transcripts may provide information that is perceived by the students as highly relevant and useful.

Even though having access to information that is relevant and useful may improve the levels of student satisfaction by the course, it is not immediately clear that such information would be promoting higher level knowledge. The results of the present study (assuming their generalizability to educational context) suggest that in discussions in which learners are left to themselves explicitation may result in better perceived information quality, but at the same time evaluation may be ineffective, and

information integrity may not be promoted by the discussions. Therefore, instructors need to be involved in the discussions, providing evaluations and criticisms as necessary, thus strengthening the evaluation dimension of knowledge construction. Moreover, instructors should promote information integrity. For instructors, it is much easier to get involved in online discussions with the learners than for health practitioners to get involved in online discussions with patients or with members of the population at risk—instructor participation is expected by the learners and does not have the potential negative consequences faced by health practitioners.

6.4 Limitations of the Study

This section discusses the limitations of the present study and suggests ways to overcome them.

As discussed at the end of the section 4.3.4, the method used in this study (quantitative content analysis of cross-sectional data) did not allow distinguishing causes from effects—the direction of causality was not tested by fitting models to data.

The indicators of the dimensions of knowledge construction, the dimensions of perceived information quality, and information integrity were rated by different pairs of coders. Therefore, the outcomes for the high-level hypotheses in this study were affected by common method bias to a relatively small extent. On the other hand, the same pair of coders rated all of the dimensions of knowledge construction. Moreover, the same pair of coders rated all of the dimensions of perceived information quality. Thus, the outcomes for low-level hypotheses may have been affected by common method bias, and some of the discriminant validity issues may have been because of common method bias.

The measures used in this study were as much as possible based on the existing measures. Nonetheless, the measure of knowledge construction involved a considerable adaptation, the measure of perceived information quality was initially developed for use in surveys rather than for content analysis, and the approach to measuring information integrity was entirely new. A research design involving testing the measures on a separate set of texts, making adjustments to address problems, and only then using the

measures for hypotheses testing would have been more robust. This, however, was not feasible because of resource constraints. The use of the measures found to be effective in this study in future research, possibly with adjustments based on the results of this study, is desirable to further confirm their validity.

The conclusions of this study, as well as the stated implications for theory and practice, rely on abductive inferences from the relationships discovered in text. Such inferences may be flawed. A direct study involving the stakeholders as participants is desirable to confirm the inferences made in this study.

This study did not investigate empirically the detailed mechanisms behind the relationships discovered. A more detailed study, involving coding for specific patterns of interactions resulting in knowledge construction and contributing to perceived information quality is desirable.

This study highlighted the difficulties involved in measuring information integrity in health support group online discussions. A more robust measure may be considered, perhaps relying on directly using the Cochrane reviews along with expert coders, with expert coders working with the full transcript of the discussions to better understand the nature of the recommendations coded for information integrity. (One has to note, though, that because health experts are time poor, a study relying on such a measure may be very difficult and expensive to implement.)

The study was limited to a single health support group and was based on discussion threads initiated over a specific and relatively short period of time (over three months). Moreover, only typical threads in terms of the number of contributors and in terms of the number of messages were included, and very small or very large discussion threads were left out (about one third of the threads were found to be “typical”). Thus, the results should be generalized with care to other health support groups or to long discussion threads. In particular, the content studied was generated in health support group online discussions devoted to weight management. The results are not necessarily generalizable to discussions of different types, such as health support group online discussions devoted to rare or immediately life threatening conditions. Even in the actual health support group where the content has been generated, the nature of

discussion dynamics may shift over time, and generalizability from the threads covered in the present study to all threads generated in the same health support group cannot be taken for granted. When applying the results of the present study, one should compare the context of the potential application to the context of the present study, and judge if the results can be applied based on the similarity of the contexts. In other words, even though the statistical generalizability of the results to a particular population is limited, analytical generalizability, as described by Yin (2009), makes the results potentially relevant rather broadly. One has to emphasize that relatively low statistical generalizability of the results is common in MIS studies—it is rarely possible to sample at random from the population of all users, and most studies are limited to particular, relatively narrow contexts (such as limited to users at a particular organization, at a particular online community and so on—limited to a setting to which the researcher has access).

6.5 Suggestions for Future Research

The results of the present study, when combined with the results of the prior studies of knowledge construction in online discussions, suggest that that high levels of involvement in online discussions and involvement over time, along with scaffolding and moderation of the discussions, result in knowledge construction in the discussions, with knowledge construction leading to higher perceived information quality (in terms of perceived usefulness and relevance) captured in the discussion transcript (see the discussion of the nomological framework in section 6.3.1.1). From the point of view of further promoting the understanding of knowledge construction in health support online discussions, it is desirable that in future research the relationships between involvement, moderation, knowledge construction, and information quality outlined in this paragraph are tested in a single model, in the context of health support group online discussions.

More broadly, the present study was the first to test quantitatively hypotheses regarding relationships between knowledge construction and other constructs conducted in a non-educational context; more quantitative studies of knowledge construction antecedents and outcomes in non-educational contexts are needed to attain a better understanding of

the knowledge construction phenomenon, particularly, in the context of inter-organizational and intra-organizational knowledge management.

The results of the present study suggest that knowledge construction is a multidimensional construct. It is desirable that further dimensions of knowledge construction are considered in future research. A potential new dimension of knowledge construction that would be highly relevant to health support online groups is emotional knowledge construction (emotional knowledge or emotional intelligence refers to "Intelligence regarding the emotions, especially in the ability to monitor one's own or others' emotions", Editors of the American Heritage Dictionaries, 2007). The process of emotional knowledge construction is likely to be distinguishable from health-related factual or procedural knowledge construction, because it relies more on empathy than on logic.

The results of the present study suggest that knowledge construction in health support group online discussions enables contributors and lurkers to access better quality information (in terms of perceived relevance and usefulness). In future research, more direct evidence of the benefits of online discussions could be obtained by measuring the effects of the discussions on health literacy of discussion contributors and of lurkers, as suggested in section 2.11. Moreover, the effects of contributors' health literacy on knowledge construction are also of interest (for a possible extension of the research model of the present study to incorporate health literacy, refer to Figure 7 in section 2.11.4).

The results of the present study suggest that knowledge construction at health support group online discussions would be promoted if moderators encourage clarifications and refinements (and, overall, promote low-level knowledge construction activities), rather than encourage arguments and evaluations (and thus promote high-level knowledge construction activities). It is desirable that in future research this proposition is tested directly, by conducting field studies.

The results of the present study suggest that knowledge construction in health support group online discussions does not contribute to information integrity. Thus, health practitioners need to be involved. As discussed in section 6.3.2.3, the most practical

way to realize such involvement is by setting up a health information portal, with a discussion forum as one of the features of the portal. Health practitioners would gain knowledge regarding the information needs and the levels of health literacy of patients and of the population at risk via the information drawn from the discussions, and would respond by making the appropriate documents based on such knowledge available via the portal, not directly participating in the discussions. It is desirable that such a portal is tested in an action research project.

In section 2.11.4 I proposed a research model that extends the research model of the present study to include health literacy as both an antecedent and an outcome of knowledge construction in health support online discussion (see Figure 7). In view of the ongoing transition to patient-centred healthcare, health literacy is of major concern in practice. Therefore, it is desirable that the extended model is tested against empirical data. This would require a multi-method approach, with content analysis conducted to measure knowledge construction (possibly, using the instruments developed in the present study) combined with surveys used to assess the levels of health literacy. A longitudinal design would be particularly appropriate, with health literacy assessed before and after health support group online discussions.

6.6 Conclusions

The ongoing transition to patient-centred health care suggests that patients take a more active role in managing their conditions. Therefore, patients need greater access to health-related information. Health support group online discussions have a potential to become an important source of health-related information because patients are directly involved in such discussions, and, therefore, the information resulting from the discussions is likely to fit the patients' levels of health literacy and cultural preferences.

Health support group online discussions can be described in terms of the process of knowledge construction—suggestions for health-related behaviours and interventions are clarified, elaborated, assessed, and justified as the discussion progresses, resulting in the creation of health-related knowledge. Both the process of knowledge creation and the resulting knowledge are reflected in the discussion transcript. The discussion transcript can be seen as a source of health-related information, and lurkers (discussion

participants who only read and never post—often the majority) access the discussion solely by reading the transcript.

The purpose of the present study was to investigate how the process of knowledge construction contributes to the quality of information in the resulting transcript. The information quality was seen as perceived information quality (the perspective of the information consumers) and information integrity (the perspective of the experts, the extent to which the information is supported by evidence-based medicine).

The study relied on content analysis to test the hypotheses that knowledge construction contributes to perceived information quality and to information integrity. The concepts were operationalized as coding schemes and independent coders were employed to code the content of a health support group discussion devoted to weight management. PLS SEM was used to test the hypotheses.

The results suggest that knowledge construction in health support group online discussions results in better perceived information quality. Lower-level knowledge construction activities—refinement, elaboration, clarification, confirmation, and negation—result in information that appears to be more relevant and more useful. Nonetheless, higher-level knowledge construction activities—critical discussion, argumentation, reasoning, and justification—do not contribute to perceived information quality.

Nonetheless, the study found no evidence that knowledge construction activities (either high- or low-level) contribute to information integrity. The results were consistent with a view that knowledge construction at health support group online discussions does not affect the extent to which the resulting information is supported by evidence-based medicine.

Based on the findings, the study made specific recommendations for managing health support group online discussions. Knowledge construction in health support group online discussions contributes to better perceived information quality. Therefore, such discussions should be promoted if the integrity of the information generated in the discussions is high enough. Discussion moderators should promote low-level

knowledge construction activities, which were found to contribute to perceived information quality.

Because the knowledge construction activities by patients contributing to health support group online discussions do not contribute to information integrity (as suggested by the findings of the present study), it is desirable that qualified health practitioners participate in the discussions. The study proposed how such participation can be achieved.

Appendix A. Knowledge Construction Form

Fill in this form for each thread.

Read carefully before filling the form.

Thread number:

Explicitation

This discussion thread involves cases of

- Weight control related **refinement** (improving or perfecting by pruning or polishing)

Strongly Disagree							**Strongly Agree**
☐	☐	☐	☐	☐	☐	☐	☐

- Weight control related **elaboration** (developing in intricate and painstaking detail)

Strongly Disagree							**Strongly Agree**
☐	☐	☐	☐	☐	☐	☐	☐

- Weight control related **clarification** (an interpretation that removes obstacles to understanding)

Strongly Disagree							**Strongly Agree**
☐	☐	☐	☐	☐	☐	☐	☐

- Weight control related **confirmation** (additional proof that something that was believed - some fact or hypothesis or theory - is correct)

Strongly Disagree							**Strongly Agree**
☐	☐	☐	☐	☐	☐	☐	☐

- Weight control related **negation** (a negative statement; a statement that is a refusal or denial of some other statement)

Strongly Disagree							**Strongly Agree**
☐	☐	☐	☐	☐	☐	☐	☐

Evaluation

This discussion thread involves cases of

- Weight control related **critical discussion** (discussion characterized by careful evaluation and judgment)

Strongly Disagree							**Strongly Agree**
☐	☐	☐	☐	☐	☐	☐	☐

- Weight control related **argumentation** (a discussion in which reasons are advanced for and against some proposition or proposal)

Strongly Disagree							**Strongly Agree**
☐	☐	☐	☐	☐	☐	☐	☐

- Weight control related **reasoning** (presentation reasons and arguments)

Strongly Disagree							**Strongly Agree**
☐	☐	☐	☐	☐	☐	☐	☐

- Weight control related **justification** (the act of defending or explaining or making excuses for by reasoning)

Strongly Disagree							**Strongly Agree**
☐	☐	☐	☐	☐	☐	☐	☐

Appendix B. Perceived Information Quality Form

Thread number:

Part I. Information Quality Performance

All items were measured on continuous 11-point semantic differential scale, *where 0 = not at all and 11 = highly.*

Please choose the answer by crossing (X) or checked (√) the right column

The information is

Relevance	0 = not at all										10 = highly
	0	1	2	3	4	5	6	7	8	9	10
Applicable to my interest in weight control											
Related to my interest in weight control											
Pertinent to my interest in weight control											
In general, information is relevant to my interest in weight control											

The information is

Understandability	0 = not at all										10 = highly
	0	1	2	3	4	5	6	7	8	9	10
Clear in meaning											
Easy to comprehend											
Easy to read											
In general, information is understandable											

The information is

Usefulness	0 = not at all										10 = highly
	0	1	2	3	4	5	6	7	8	9	10
Informative to making decisions regarding approaches to weight control											
Valuable to making decisions regarding approaches to weight control											
In general, information that is useful											

Appendix C. Information Integrity Data Extraction Form

Introduction

Please, read each thread carefully, and note any recommendations for action relevant to weigh control that are explicitly or implicitly suggested by the discussion. (Just ask yourself: Which actions relevant to weight control one would consider as viable solely on the basis of reading this particular thread?)

In some threads there could more than one recommendation. In other threads, there could be no recommendations at all.

Before executing the following steps, read the whole of the thread.

Steps

1. **Highlight**

For each recommendation, highlight one occurrence in the text of the thread (the first occurrence in the thread, or the occurrence in the thread where the recommendation is given most clearly).

Example 1

Bob: A combination of physical exercise with diet is effective, if you have the perseverance to keep up the regime.

Ann: I agree. Tough, but effective. My doctor told me this would work, for as long as I do not overdo it.

Peter: A friend of mine tried this. It did work.

2. **Clarify**

When the way the recommendation is given in the text of the thread is not self-explanatory enough (e.g., in case of implicit recommendations), clarify the meaning of the recommendation you found by writing on the margin.

Example 2

Need to reduce alcohol to control weight

Bob: Oh, no, I've been drinking too much beer after the game on Friday. This is taking me nowhere …

Ann: Yeh, Bob, you should be careful.

Peter: Does it matter? There is no fat in beer. I think it matters what you eat, not what you drink.

Appendix D. Information Integrity Expert Assessment

All items were measured on continuous 8-point semantic differential scale, where 0= highly anecdotal and 7 = highly evidence based.

Please choose the answer by crossing (x) or tick (v) in the right column.

Mental State

a.	For people who interested in weight loss, *positive mood* will help contribute to their weight loss program.	<<Highly Anecdotal							Highly Evidence Based>>
		0	1	2	3	4	5	6	7
b.	For people who interested in weight loss, *mental state* can affect their weight loss journey.	<<Highly Anecdotal							Highly Evidence Based>>
		0	1	2	3	4	5	6	7
c.	People who interested in weight loss should try to change things around for the variety of their weight loss program.	<<Highly Anecdotal							Highly Evidence Based>>
		0	1	2	3	4	5	6	7
d.	People who interested in weight loss need *mentally prepared* for weight loss journey.	<<Highly Anecdotal							Highly Evidence Based>>
		0	1	2	3	4	5	6	7
e.	People who interested in weight loss should *stay focus* on their goal for weight loss.	<<Highly Anecdotal							Highly Evidence Based>>
		0	1	2	3	4	5	6	7
f.	Way of thinking – mental attitude = weight control. *Change way of thinking and accept this isn't short term of diet but their new way of life.*	<<Highly Anecdotal							Highly Evidence Based>>
		0	1	2	3	4	5	6	7
g.	Weather contributes in weight loss journey by giving positive vibes.	<<Highly Anecdotal							Highly Evidence Based>>
		0	1	2	3	4	5	6	7

h.	People who interested in weight loss should not go on the scales constantly if they are stress because of it.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
i.	For people who interested in weight loss should avoid too much eat and drink when they are feeling down.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Water

a.	For people who interested in weight loss, they must drink more water in their weight loss journey.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	Drink water helps people who interested in weight loss and they should follow suggested drinking quota.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
c.	For people who interested in weight loss should drink 2 to 3 litres of water every day.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
d.	For people who interested in weight loss, they could put water available near them for drinking, so they don't have to eat any food when they are hungry.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
e.	Drinking water is key to success for weight loss.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

f.	Drinking eight glass of water every day increased fat burning, healthier skin, better digestion, and fewer cravings.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
g.	Drinking water is the key for weight loss and the suggestion is to drink 2 litres ++ every day.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
h.	Hot drinks can help as well like herbal teas – helps to make us feel full.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
i.	Drinking water is major factor for weight loss journey.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Exercise

a.	For people who interested in weight loss, heaps variation of exercise will contribute to their weight loss program.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	Intensity of exercise may affect weight loss.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
c.	Exercise helps us live longer.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

d.	Exercise effective for weight control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
e.	Treadmill for a half hour everyday helps for weight loss.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
f.	A small amount of exercise is better than none.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
g.	Exercise is very important in weight loss journey.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
h.	Exercise for post breast cancer: 20 minutes walking for 5 times a week.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
i.	Up activity levels to boost up weight loss. Do longer or more intense or both workouts.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
j.	Do exercise that is fun and enjoyable, like zumba for example. Zumba is the name of dance fitness.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
k.	For people who interested in weight loss, variation of exercises will help them in weight loss journey.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
l.	Night time exercise will be solutions for someone that busy during the daytime.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

m.	For people who interested in weight loss and has problem with one paralyzed leg should try alternative of exercises, such as Pilates and swimming.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
n.	Routine exercise does help for weight loss.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
o.	For people who interested in weight loss should make their gym gear available at home, so they could exercise every day.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Vegetable soups

a.	For people who interested in weight loss, Vegetable soups will sustains their hunger.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	A vegetable soup is healthy and keeps us full and helps weight control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Food variation

a.	For people who interested in weight loss, eating a good variety of foods from all the food groups will helps their weight control program.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	Change and variety in foods helps move weight.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
c.	For people who interested in weight loss should try different/variety of foods for weight control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
d.	Variation for breakfast like fruit and yoghurt will help us control our weight.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
e.	For people who interested in weight loss should choose the lower fat variety for treats.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Healthy and fresh food

a.	For people who interested in weight loss, they have to plan their diet with wholesome healthy balance food.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

b.	For people who interested in weight loss, they should make sure they have fresh and healthy food in their 'weight loss' menus.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
c.	For people who interested in weight loss should cut down on processed food and include protein in every meal and snack.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
d.	Fresh food that isn't full of preservatives and additives is a first stop for stocking our fridge. Healthy and fresh food = weight control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Motivation

a.	For people who interested in weight loss, join weight loss community and get support from others will keep them motivated.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	Be inspired by others, will help them (people who interested in weight loss) motivated.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
c.	Question our self is good motivation in weight loss journey.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
d.	Supports from friends help with weight control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

e.	Set our goals will keep us motivated in our weight loss journey.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
f.	Challenge ourselves to get up and start moving. We have to make a big effort for weight loss.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
g.	Makes short term and long term goal to focus on.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
h.	Start slow but always keep in track.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
i.	It is not just about eating well and exercising it is the psychological causes, we can sometimes overlook that make easy for weight to creep back in.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
j.	Being accountable by sharing our weight loss journey with others helps with weight control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
k.	Weight loss journey will always be work in progress.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
l.	Realizing and accepting about overweight is part of weight control, and try to get support from partner, family and friends.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
m.	Do not be discouraged if weight loss doesn't meet expectation.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

n.	For people who interested in weight loss need to be optimistic rather than pessimistic in their journey.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Plan and weight menu

a.	Plan effectively: writing or tracking everything we eat is effective in helping with weight loss.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	Plan weekly menus is effective for people who interested in weight loss.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
c.	For people who interested in weight loss should take lunch to work to keep with their weight loss plan.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
d.	For people who interested in weight loss, they should calculating points at each meal to follow their weight loss program.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
e.	Having food prepared for the day to make sure we have food according to weight loss program.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
f.	Weighing food according to weight loss program will contributes to weight control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

g.	For people who interested in weight loss should use eating guide as their reference for preparing the food.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
h.	Plan snacks. Snacks can be a huge downfall if people who interested in weight loss don't identify and keep good stock of it and take it with them when they go out.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
i.	For take away options make sure the food is healthy.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
j.	Good eating habits contribute to weight loss and people who interested in weight loss have to watch the serving size.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
k.	Break up our main meals into smaller more often meals so that we always have something to eat and keep lots of low points but bulky/long lasting food available.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Suggestions for weight loss

a.	Drinking water, tracking our food, and exercise are obviously the key to weight loss.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

b.	Positive thinking and exercise will contribute to weight loss program.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
c.	Strict with our points, drank plenty of water, and exercised 3 times a day.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
d.	Eating healthy is a life plan, having support and accountability helps with weight control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
e.	Stay positive; drink plenty of water, get outside, and do some exercise.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Medical condition

a.	Medical conditions may affect weight loss/control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	For people who interested in weight loss and has special conditions should ask their doctor about their weight loss program	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

c.	After being sick if able eat light meal, high in protein and carbohydrate.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Protein

a.	More protein can make us feel fuller and eat less (weight control).	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	For people who interested in weight loss, they should increase their protein intake with lean protein such as fish, chicken, lean meats, tofu, baked beans, etc.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
c.	For people who interested in weight loss, they should eat protein with each meal.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Supplements

a.	For people who interested in weight loss should change their snacks with fruits.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	For people who interested in weight loss should try taking olive leaf extract available from health food stores (Vitamins).	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
c.	For people who interested in weight loss, they should eat a variety of fruit and veggies every day for the most nutritional benefits.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
d.	For people who interested in weight loss, they should try to change their breakfast with banana, honey, and toast for the supplement in the morning.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Carbohydrates

a.	For people who interested in weight loss, they should knock out refined carbohydrates.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	Carbohydrates make it difficult to control weight.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

c.	Carbohydrates are one of the main types of food. Our live breaks down carbohydrates into glucose (blood sugar). Our body uses this sugar for energy for your cells, tissues, and organs.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
d.	For people who interested in weight loss, they should try swapping some of their carbohydrates for increased protein portions.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
e.	Reduced substantially sugar, fat, and sodium intake for weight loss journey.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
f.	Cut out sugar and artificial sweeteners.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Medical advisor

a.	When pregnant seek medical advice for weight control.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	Seeing a dietician will be help to reinforce all those good habits in weight loss journey.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Suggestions

a.	Wholegrain better than starchy foods.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7
b.	For people who interested in weight loss, they should eat their dinner after their exercise to make them full in the night.	<<Highly Anecdotal Highly Evidence Based>> 0 1 2 3 4 5 6 7

Note:

……

……

……

Appendix E. Details of Post-Hoc Analyses

A number of post-hoc analyses were conducted to better understand the meaning of the initial results.

E.1 A Model With Both Knowledge Construction and Perceived Information Quality Treated as One-Dimensional Constructs—The Details of Measurement Model Analysis

This section describes the results of measurement model analysis for the post-hoc analysis described in section 5.8.1.

Item reliability. All factor loadings of the items on their own constructs were above the threshold value of 0.7 (see Table E-1).

Table E-1 Indicator Loadings

	Knowledge construction	Perceived Information Quality	Information Integrity
Knowledge construction			
Refinement	0.786	0.275	-0.034
Elaboration	0.763	0.210	0.038
Clarification	0.835	0.258	-0.002
Confirmation	0.833	0.265	-0.080
Negation	0.713	0.251	0.114
Critical discussion	0.881	0.213	0.055
Argumentation	0.887	0.196	0.078
Reasoning	0.894	0.208	0.072
Justification	0.820	0.212	0.012
Perceived information quality			
Applicable	0.271	0.983	-0.051
Related	0.295	0.979	-0.030
Pertinent	0.284	0.984	-0.066
Relevant in general	0.289	0.985	-0.054
Informative	0.258	0.952	-0.128
Valuable	0.277	0.975	-0.117
Useful in general	0.280	0.974	-0.093
Information integrity			
Average	0.031	-0.078	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Internal consistency reliability. All of the Cronbach's $\alpha$and composite reliability (ρ) values were considerably higher than the threshold value of 0.7 (see Table E-2).

Thus, in terms of item reliability and internal consistency reliability, the measures were reliable.

Table E-2 Internal Consistency Indices

	Cronbach's α	Composite reliability (ρ)
Knowledge construction	0.941	0.950
Perceived information quality	0.992	0.993
Information integrity	1.000	1.000a

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Convergent validity. All of the AVE values were above the threshold of 0.5 (see Table E-3). This, along with the evidence of item and internal consistency reliability, suggested that the multi-indicator constructs had convergent validity.

Table E-3 Average Variance Extracted

Construct	AVE
Knowledge construction	0.681
Perceived information quality	0.953
Information Integrity	1.000a

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Discriminant validity. Items loaded on their own constructs higher than on any other constructs in the model, suggesting discriminant validity (see Table E-1). Moreover, for all of the constructs, the square root of AVE was larger than the correlations of the construct with other constructs in the model (see Table E-4). The absolute values of correlations between latent constructs were considerably lower than the threshold value of 0.7. Thus, there were no discriminant validity issues.

Table E-4 Correlations Between Latent Variables

	Knowledge construction	Perceived information quality	Information Integrity
Knowledge construction	0.825		
Perceived information integrity	0.286	0.976	
Information Integrity	0.031	-0.087	1.000

Note. The values on the diagonal are square roots of AVE.

E.2 A Model With Explicitation and Evaluation as Separate Constructs and With Perceived Information Quality Treated as a One-Dimensional Construct—The Details of Measurement Model Analysis

This section describes the results of measurement model analysis for the post-hoc analysis described in section 5.8.2.

Item reliability. All factor loadings were above the threshold value of 0.7; the only exception (as in the analysis in section 5.6) was the negation indicator of explicitation, which was very close to 0.7 (see Table E-5).

Table E-5 Indicator Loadings on the Dimensions of Knowledge Construction, Perceived Information Quality, and Information Integrity

	Explicitation	Evaluation	Perceived Information Quality	Information Integrity
Explicitation				
Refinement	0.863	0.563	0.275	-0.034
Elaboration	0.847	0.534	0.210	0.038
Clarification	0.863	0.667	0.258	-0.002
Confirmation	0.893	0.626	0.265	-0.080
Negation	0.692	0.619	0.251	0.114
Evaluation				
Critical discussion	0.704	0.969	0.212	0.055
Argumentation	0.705	0.979	0.196	0.078
Reasoning	0.714	0.982	0.207	0.072
Justification	0.654	0.898	0.212	0.012
Perceived information quality				
Applicable	0.289	0.204	0.982	-0.051
Related	0.305	0.235	0.978	-0.030
Pertinent	0.293	0.226	0.983	-0.066
Relevant in general	0.299	0.230	0.984	-0.054
Informative	0.287	0.178	0.954	-0.128
Valuable	0.305	0.197	0.976	-0.117
Useful in general	0.305	0.202	0.975	-0.093
Information integrity				
Average	0.006	0.057	-0.079	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Internal consistency reliability. All of the Cronbach's α and composite reliability (ρ) values were considerably higher than the threshold value of 0.7 (see Table E-6).

Therefore, in terms of item reliability and internal consistency reliability, the measures were reliable.

Table E-6 Internal Consistency Indices

	Cronbach's α	Composite reliability (ρ)
Explicitation	0.889	0.919
Evaluation	0.969	0.978
Perceived information quality	0.992	0.993
Information integrity	1.000[a]	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Convergent validity. All of the AVE values were above the threshold of 0.5 (see Table E-7). This, along with the evidence of item and internal consistency reliability, suggested that the multi-indicator constructs had convergent validity.

Table E-7 Average Variance Extracted

Construct	AVE
Explicitation	0.697
Evaluation	0.917
Perceived information quality	0.953
Information Integrity	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Discriminant validity. Items loaded on their own constructs higher than on any other constructs in the model, suggesting discriminant validity(see Table E-5). Moreover, for all of the constructs, the square root of AVE was larger than the correlations of the construct with other constructs in the model (see Table E-8). The correlation between

explicitation and evaluation was marginally higher than the threshold value of 0.7 (as in the analysis in section 5.6); the rest of the correlations between constructs were clearly below 0.7. Thus, there were no major issues with discriminant validity.

Table E-8 Latent Variables Correlation Matrix

	Explicitation	Evaluation	Perceived information quality	Information Integrity
Explicitation	0.835			
Evaluation	0.726	0.958		
Perceived information quality	0.305	0.216	0.976	
Information Integrity	0.006	0.057	-0.050	1.000

Note. The values on the diagonal are square roots of AVE.

E.3 A Model With Explicitation Hypothesized to Affect Evaluation—The Details of Measurement Model Analysis

This section describes the results of measurement model analysis for the post-hoc analysis described in section 5.8.3.

The model included the same constructs as in the model discussed in section 5.6, measured by the same items. The model was fitted to the same data set. Nonetheless, because the properties of a measurement model in PLS depend on the relationships between constructs in the model, the measurement model was re-examined.

Item reliability. All factor loadings were above the threshold value of 0.7; the only exception (as in the analysis in section 5.6) was the negation indicator of explicitation, which was extremely close to 0.7 (see Table E-9).

Table E-9 Indicator Loadings on the Dimensions of Knowledge Construction, Perceived Information Quality, and Information Integrity

	Explicitation	Evaluation	Relevance	Usefulness	Information integrity
Explicitation					
Refinement	0.855	0.563	0.278	0.259	-0.034
Elaboration	0.841	0.534	0.213	0.197	0.038
Clarification	0.869	0.667	0.246	0.264	-0.002
Confirmation	0.894	0.626	0.255	0.267	-0.080
Negation	0.697	0.618	0.239	0.256	0.114
Evaluation					
Critical discussion	0.706	0.969	0.219	0.194	0.055
Argumentation	0.707	0.980	0.202	0.179	0.078
Reasoning	0.716	0.983	0.216	0.187	0.072
Justification	0.657	0.897	0.226	0.184	0.012
Relevance					
Applicable	0.288	0.204	0.995	0.924	-0.051
Related	0.304	0.235	0.995	0.914	-0.030
Pertinent	0.292	0.226	0.994	0.927	-0.066
Relevant in general	0.298	0.230	0.998	0.925	-0.054
Usefulness					
Informative	0.287	0.178	0.899	0.987	-0.128
Valuable	0.304	0.197	0.931	0.994	-0.117
Useful in general	0.305	0.202	0.929	0.996	-0.093
Information integrity					
Average	0.007	0.057	-0.050	-0.113	1.000

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Internal consistency reliability. All of the Cronbach's α and composite reliability (ρ) values were considerably higher than the threshold value of 0.7 (see Table E-10).

Therefore, in terms of item reliability and internal consistency reliability, the measures were reliable.

Table E-10 Internal Consistency Indices

	Cronbach's α	Composite reliability (ρ)
Explicitation	0.889	0.919
Evaluation	0.969	0.978
Relevance	0.997	0.998
Usefulness	0.992	0.995
Information Integrity	1.000[a]	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Convergent validity. All of the AVE values were above the threshold of 0.5 (see Table E-11). This, along with the evidence of item and internal consistency reliability, suggested that the multi-indicator constructs had convergent validity.

Table E-11 Average Variance Extracted

Construct	AVE
Explicitation	0.696
Evaluation	0.917
Relevance	0.991
Usefulness	0.985
Information Integrity	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Discriminant validity. To assess discriminant validity, loadings of items on their own constructs were compared with their loading on other constructs in the model (see Table E-9). In all cases, items loaded on their own constructs higher than on any other constructs in the model, suggesting discriminant validity.

The negation indicator of explicitation loaded on evaluation almost as high as on its own construct. The correlation between relevance and usefulness was very high (0.992), and the correlation between explicitation and evaluation was rather high (0.728); both of these correlations were higher than the threshold value of 0.7. Thus, the issues with discriminant validity were the same as in the model with no relationship between explicitation and evaluation.

For all of the constructs, the square root of AVE was larger than the correlations of the construct with other constructs in the model (see Table E-12).

Table E-12 Latent Variables Correlation Matrix

	Explicitation	Evaluation	Relevance	Usefulness	Information Integrity
Explicitation	0.834				
Evaluation	0.728	0.958			
Relevance	0.297	0.225	0.995		
Usefulness	0.301	0.194	0.927	0.992	
Information Integrity	0.005	0.057	-0.050	-0.113	1.000

Note. The values on the diagonal are square roots of AVE.

E.4 A Model of Coder's Conceptions—The Details of Measurement Model Analysis

This section describes the results of measurement model analysis for the post-hoc analysis described in section 5.8.4.

Item reliability. To assess item reliability, the loadings of items on their own constructs were examined (see Table E-13, Table E-14, Table E-15, Table E-16, and Table E-17). In almost all cases, the loadings were higher than the threshold value of 0.7; the only exception was negation, which was still either higher than 0.6 or (in case of the coder 2 conceptions) less than 0.6, but very close to it (0.585).

Thus, overall, item reliability was judged to be acceptable.

Table E-13 Indicator Loadings on Explicitation and Its Coder's Conceptions

	Explicitation	Explicitation by coder 1	Explicitation by coder 2	Evaluation	Evaluation by coder 1	Evaluation by coder 2	Relevance	Relevance by coder 3	Relevance by coder 4	Usefulness	Usefulness by coder 3	Usefulness coder 4	Information integrity	Information integrity by expert coder 1	Information integrity by expert coder 2
Explicitation															
Refinement	0.880	0.808	0.815	0.563	0.545	0.506	0.278	0 269	0.224	0.259	0.241	0.224	-0.034	0.002	-0.066
Elaboration	0.869	0.821	0.781	0.534	0.526	0.468	0.213	0.235	0.148	0.197	0.204	0.151	0.038	0.088	-0.036
Clarification	0.858	0.771	0.793	0.668	0.648	0.599	0.246	0.257	0.183	0.264	0.291	0.184	-0.002	0.028	-0.038
Confirmation	0.893	0.816	0.813	0.626	0.626	0.536	0.255	0.238	0.213	0.267	0.262	0.218	-0.080	0.008	-0.159
Negation	0.655	0.641	0.529	0.618	0.619	0.526	0.239	0.238	0.188	0.256	0.267	0.194	0.114	0.143	0.039
Explicitation by coder 1															
Refinement 1	0.799	0.844	0.619	0.532	0.555	0.425	0.289	0.281	0.232	0.279	0.268	0.233	-0.056	-0.004	-0.101
Elaboration 1	0.784	0.854	0.578	0.483	0.515	0.370	0.223	0.235	0.164	0.215	0.214	0.172	0.014	0.079	-0.071
Clarification 1	0.750	0.799	0.549	0.630	0.664	0.494	0.235	0.268	0.159	0.278	0.330	0.172	0.028	0.078	-0.042
Confirmation 1	0.763	0.836	0.535	0.591	0.651	0.426	0.212	0.207	0.170	0.235	0.244	0.178	-0.062	0.039	-0.164
Negation 1	0.606	0.666	0.403	0.568	0.606	0.432	0.231	0.233	0.179	0.251	0.255	0.195	0.098	0.122	0.036
Explicitation by coder 2															
Refinement 2	0.840	0.641	0.924	0.512	0.449	0.525	0.223	0.213	0.181	0.195	0.171	0.178	-0.003	0.008	-0.015
Elaboration 2	0.831	0.646	0.905	0.510	0.453	0.513	0.165	0.197	0.105	0.143	0.159	0.100	0.062	0.085	0.012
Clarification 2	0.804	0.580	0.909	0.573	0.496	0.596	0.207	0.192	0.173	0.193	0.187	0.159	-0.037	-0.036	-0.026
Confirmation 2	0.834	0.608	0.938	0.522	0.457	0.537	0.245	0.218	0.213	0.243	0.221	0.214	-0.082	-0.030	-0.118
Negation 2	0.535	0.396	0.585	0.511	0.447	0.525	0.179	0.172	0.145	0.186	0.204	0.132	0.105	0.135	0.033

Table E-14 Indicator Loadings on Evaluation and Its Coder's Conceptions

	Explicitation	Explicitation by coder 1	Explicitation by coder 2	Evaluation	Evaluation by coder 1	Evaluation by coder 2	Relevance	Relevance by coder 3	Relevance by coder 4	Usefulness	Usefulness by coder 3	Usefulness coder 4	Information integrity	Information integrity by expert coder 1	Information integrity by expert coder 2
Evaluation															
Critical discussion	0.695	0.666	0.596	0.969	0.925	0.891	0.219	0.288	0.117	0.193	0.245	0.105	0.055	0.100	-0.019
Argumentation	0.696	0.672	0.594	0.980	0.933	0.905	0.202	0.247	0.123	0.179	0.213	0.111	0.078	0.103	0.021
Reasoning	0.705	0.678	0.603	0.983	0.933	0.912	0.215	0.264	0.131	0.187	0.219	0.119	0.072	0.106	0.006
Justification	0.641	0.642	0.516	0.896	0.868	0.799	0.225	0.272	0.140	0.184	0.215	0.118	0.012	0.034	-0.019
Evaluation by coder 1															
Critical discussion 1	0.680	0.704	0.524	0.927	0.960	0.751	0.217	0.309	0.098	0.186	0.244	0.093	0.063	0.119	-0.027
Argumentation 1	0.690	0.720	0.525	0.942	0.981	0.756	0.213	0.263	0.126	0.191	0.224	0.121	0.088	0.124	0.013
Reasoning 1	0.693	0.726	0.525	0.945	0.985	0.759	0.202	0.267	0.106	0.172	0.213	0.099	0.076	0.116	0.001
Justification 1	0.612	0.665	0.431	0.814	0.875	0.609	0.200	0.261	0.109	0.166	0.201	0.100	-0.001	0.032	-0.042
Evaluation by coder 2															
Critical discussion 2	0.615	0.516	0.612	0.887	0.741	0.960	0.190	0.217	0.128	0.176	0.210	0.108	0.035	0.057	-0.004
Argumentation 2	0.604	0.514	0.596	0.889	0.738	0.966	0.160	0.190	0.101	0.137	0.167	0.082	0.054	0.061	0.028
Reasoning 2	0.614	0.515	0.613	0.884	0.726	0.973	0.201	0.220	0.142	0.178	0.195	0.125	0.057	0.077	0.012
Justification 2	0.548	0.474	0.527	0.820	0.674	0.900	0.214	0.230	0.154	0.170	0.189	0.118	0.028	0.028	0.018

Table E-15 Indicator Loadings on Relevance and Its Coder's Conceptions

	Explicitation	Explicitation by coder 1	Explicitation by coder 2	Evaluation	Evaluation by coder 1	Evaluation by coder 2	Relevance	Relevance by coder 3	Relevance by coder 4	Usefulness	Usefulness by coder 3	Usefulness coder 4	Information integrity	Information integrity by expert coder 1	Information integrity by expert coder 2
Relevance															
Applicable	0.285	0.288	0.228	0.204	0 198	0.182	0.995	0.828	0.906	0.924	0.767	0.885	-0.051	-0.098	0.024
Related	0.301	0.299	0.248	0.235	0.227	0.212	0.995	0.850	0.889	0.914	0.770	0.865	-0.030	-0.079	0.040
Pertinent	0.289	0.292	0.230	0.226	0.221	0.200	0.994	0.826	0.906	0.927	0.774	0.885	-0.066	-0.109	0.009
Relevant in general	0.295	0.300	0.233	0.230	0.224	0.205	0.998	0.835	0.905	0.924	0.770	0.884	-0.054	-0.097	0.017
Relevance by coder 3															
Applicable 3	0.296	0.305	0.232	0.274	0 284	0.220	0.833	0.993	0.525	0.714	0.765	0.521	0.095	0.084	0.076
Related 3	0.288	0.292	0.230	0.275	0.281	0.225	0.825	0.992	0.513	0.697	0.753	0.502	0.142	0.127	0.111
Pertinent 3	0.300	0.311	0.230	0.285	0.295	0.230	0.839	0.992	0.536	0.727	0.776	0.533	0.122	0 113	0.092
Relevant in general 3	0.290	0.302	0.222	0.276	0.288	0.219	0.837	0.998	0.527	0.713	0.758	0.525	0.099	0.091	0.074
Relevance by coder 4															
Applicable 4	0.213	0.211	0.174	0.105	0.088	0.112	0.894	0.517	0.990	0.875	0.597	0.961	-0.150	-0.213	-0.020
Related 4	0.244	0.236	0.206	0.149	0.131	0.152	0.907	0.542	0.991	0.882	0.610	0.963	-0.163	-0.228	-0.027
Pertinent 4	0.215	0.211	0.178	0.130	0.114	0.132	0.891	0.513	0.989	0.874	0.598	0.959	-0.198	-0.257	-0.057
Relevant in general 4	0.233	0.232	0.190	0.144	0.126	0.148	0.901	0.525	0.996	0.882	0.609	0.964	-0.160	-0.220	-0.031

Table E-16 Indicator Loadings on Usefulness and Its Coder's Conceptions

	Explicitation	Explicitation by coder 1	Explicitation by coder 2	Evaluation	Evaluation by coder 1	Evaluation by coder 2	Relevance	Relevance by coder 3	Relevance by coder 4	Usefulness	Usefulness by coder 3	Usefulness coder 4	Information integrity	Information integrity by expert coder 1	Information integrity by expert coder 2
Usefulness															
Informative	0.283	0.308	0.199	0.178	0.177	0.155	0.899	0.687	0.866	0.987	0.886	0.881	-0.128	-0.125	-0.087
Valuable	0.299	0.310	0.229	0.197	0.190	0.179	0.931	0.727	0.885	0.994	0.881	0.900	-0.117	-0.115	-0.078
Useful in general	0.299	0.310	0.229	0.202	0.194	0.184	0.929	0.721	0.886	0.996	0.884	0.900	-0.093	-0.102	-0.050
Usefulness by coder 3															
Informative 3	0.266	0.300	0.176	0.197	0.197	0.167	0.707	0.696	0.560	0.837	0.964	0.544	-0.052	0.018	-0.120
Valuable 3	0.298	0.315	0.223	0.234	0.233	0.202	0.780	0.782	0.607	0.886	0.983	0.612	-0.020	0.040	-0.086
Useful in general 3	0.316	0.339	0.229	0.251	0 250	0.217	0.785	0.775	0.620	0.893	0.991	0.618	0.014	0.062	-0.050
Usefulness by coder 4															
Informative 4	0.229	0.241	0.172	0.117	0.114	0.106	0.863	0.510	0.947	0.889	0.596	0.988	-0.168	-0.230	-0.034
Valuable 4	0.236	0.240	0.186	0.119	0.107	0.118	0.883	0.519	0.972	0.892	0.595	0.994	-0.188	-0.244	-0.055
Useful in general 4	0.224	0.221	0.184	0.115	0.103	0.115	0.886	0.530	0.969	0.901	0.610	0.996	-0.177	-0.239	-0.041

Table E-17 Indicator Loadings on Perceived Information Quality and Its Coder's Conceptions

	Explicitation	Explicitation by coder 1	Explicitation by coder 2	Evaluation	Evaluation by coder 1	Evaluation by coder 2	Relevance	Relevance by coder 3	Relevance by coder 4	Usefulness	Usefulness by coder 3	Usefulness coder 4	Information integrity	Information integrity by expert coder 1	Information integrity by expert coder 2
Information integrity															
Average	0.001	0.000	0.002	0.058	0.061	0.046	-0.051	0.115	-0.169	-0.113	-0.019	-0.179	1.000[a]	0.875	0.806
Information integrity by expert coder 1															
Average 1	0.058	0.074	0.029	0.090	0.105	0.059	-0.096	0.104	-0.231	-0.115	0.041	-0.240	0.875	1.000[a]	0.419
Information integrity by expert coder 2															
Average 2	-0.068	-0.091	-0.031	-0.002	-0.013	0.014	0.022	0.089	-0.034	-0.072	-0.086	-0.043	0.806	0.419	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Internal consistency reliability. All of the Cronbach's α and composite reliability (ρ) values were considerably higher than the threshold value of 0.7 (see Table E-18).

Therefore, in terms of internal consistency reliability the measures were reliable.

Table E-18 Internal Consistency Indices

	Cronbach's α	Composite reliability (ρ)
Explicitation	0.889	0.920
Explicitation by coder 1	0.860	0.900
Explicitation by coder 2	0.907	0.934
Evaluation	0.969	0.978
Evaluation by coder 1	0.964	0.974
Evaluation by coder 2	0.964	0.974
Relevance	0.997	0.998
Relevance by coder 3	0.996	0.997
Relevance by coder 4	0.994	0.996
Usefulness	0.992	0.995
Usefulness by coder 3	0.979	0.986
Usefulness by coder 4	0.993	0.995
Information Integrity	1.000[a]	1.000[a]
Information integrity by expert coder 1	1.000[a]	1.000[a]
Information integrity by expert coder 2	1.000[a]	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Convergent validity. All of the AVE values were above the threshold of 0.5 (see Table E-19). This, along with the evidence of item and internal consistency reliability, suggested that the multi-indicator constructs had convergent validity.

Table E-19 Average Variance Extracted

Construct	AVE
Explicitation	0.698
Explicitation by coder 1	0.645
Explicitation by coder 2	0.744
Evaluation	0.917
Evaluation by coder 1	0.905
Evaluation by coder 2	0.902
Relevance	0.991
Relevance by coder 3	0.988
Relevance by coder 4	0.983
Usefulness	0.985
Usefulness by coder 3	0.959
Usefulness by coder 4	0.985
Information Integrity	1.000[a]
Information integrity by expert coder 1	1.000[a]
Information integrity by expert coder 2	1.000[a]

[a]This value was exactly one, and trivial, because information integrity had only one indicator.

Discriminant validity. To assess discriminant validity, loadings of items on their own constructs were compared with their loading on other constructs in the model (see Table E-13, Table E-14, Table E-15, Table E-16, and Table E-17). For constructs, items loaded higher on their own constructs than on other constructs.

As a further assessment of discriminant validity, the square root of AVE for each latent construct was compared to correlations of the construct with other constructs (see Table E-20). According to this criterion, the model did not discriminate sufficiently between explicitation and explicitation by coder 1, explicitation and explicitation by coder 2, and between evaluation and evaluation by coder 1. Thus, the discriminant validity issues were between constructs and their generalizations (immediately related concepts at higher level in the hierarchical model). In the spirit of the approach to hierarchical

PLS modelling suggested by Wetzels et al. (2009), I proceeded to fit the structural model without making adjustments to resolve discriminant validity issues between such constructs.

Table E-20 Latent Variables Correlation Matrix

	Explicitation	Explicitation by coder 1	Explicitation by coder 2	Evaluation	Evaluation by coder 1	Evaluation by coder 2	Relevance	Relevance by coder 3	Relevance by coder 4	Usefulness	Usefulness by coder 3	Usefulness by coder 4	Information integrity	Information integrity by expert coder 1	Information integrity by expert coder 2
Explicitation	0.836														
by coder 1	0.927	0.803													
by coder 2	0.902	0.674	0.863												
Evaluation	0.715	0.694	0.604	0.958											
by coder 1	0.704	0.740	0.529	0.956	0.951										
by coder 2	0.627	0.532	0.619	0.917	0.758	0.950									
Relevance	0.294	0.296	0.236	0.225	0.218	0.201	0.995								
by coder 3	0.295	0.305	0.230	0.279	0.289	0.225	0.839	0.994							
by coder 4	0.228	0.225	0.189	0.133	0.116	0.138	0.906	0.529	0.992						
Usefulness	0.296	0.312	0.220	0.194	0.188	0.174	0.927	0.717	0.886	0.992					
by coder 3	0.300	0.325	0.215	0.233	0.232	0.200	0.774	0.768	0.609	0.891	0.979				
by coder 4	0.232	0.236	0.182	0.118	0.109	0.114	0.884	0.524	0.970	0.901	0.605	0.993			
Information integrity	0.001	0.000	0.002	0.058	0.061	0.046	-0.051	0.115	-0.169	-0.113	-0.019	-0.179	1.000		
by expert coder 1	0.058	0.074	0.029	0.090	0.105	0.059	-0.096	0.104	-0.231	-0.115	0.041	-0.240	0.875	1.000	
by expert coder 2	-0.068	-0.091	-0.031	-0.002	-0.013	0.014	0.022	0.089	-0.034	-0.072	-0.086	-0.043	0.806	0.419	1.000

Note. The values on the diagonal are square roots of AVE.

Appendix F. Mplus Code for the EFA Analysis Discussed in Section 5.5, With the Number of Factors Set to Four

```
TITLE: Exploratory Factor Analysis

  DATA: FILE IS EFA.csv;

  VARIABLE:

      NAMES ARE exre exel excl exco exne

           evcd evar ever evju

           rlvapp rlvrltd rlvpert rlvig

           usinf usval usig;

ANALYSIS: TYPE = EFA 4 4;
```

Appendix G. Mplus Code for Linear Regression Analysis Discussed in Section 5.8.5

```
TITLE: Information integrity

  DATA: FILE IS ii.csv;

  VARIABLE:

      NAMES ARE ii exp eva;

  MODEL:

ii on exp;

ii on eva;

OUTPUT: STANDARDIZED;
```

References

Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs of American Medical Association. (1999). Health literacy: Report of the Council on Scientific Affairs. *Journal of the American Medical Association, 281*(6), 552–557.

Aguilera, C., Perez, M. A., & Alonso Palacio, L. M. (2010). Readability of diabetes education materials: Implications for reaching patients with written materials. *Revista Salud Uninorte, 26*(1), 12-26.

Alavi, M. (1994). Computer-mediated collaborative learning: An empirical evaluation. *MIS Quarterly, 18*(2), 159-174.

American Psychological Association. (2009). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Arbuckle, J. L. (2010). *IBM SPSS AMOS 19 user's guide*. Crawfordville, FL: Amos Development Corporation.

Baker, D. W. (2006). The meaning and the measure of health literacy. *Journal of General Internal Medicine, 21*(8), 878-883.

Balint, E. (1969). The possibilities of patient-centered medicine. *The Journal of the Royal College of General Practitioners, 17*(82), 269-276.

Bedi, R. P., & Alexander, D. A. (2009). Using multivariate concept-mapping for examining client understandings of counselling. *Canadian Journal of Counselling and Psychotherapy, 43*(2), 76-91.

Betancourt, J. R., Green, A. R., Carrillo, J. E., & Ananeh-Firempong, O. (2003). Defining cultural competence: A practical framework for addressing racial/ethnic disparities in health and health care. *Public Health Reports, 118*(4), 293-302.

Boland, R. J. (1985). Phenomenology: A preferred approach to research on information systems. In E. Mumford, R. Hirschheim, G. Fitzgerald, & A. T. Wood-Harper (Eds.), *Trends in information systems* (pp. 341-349). Amsterdam, Netherlands: North-Holland.

Brookes, I. (Ed.). (2003). *The Chambers Dictionary*. Edinburgh, United Kingdom: Chambers Harrap.

Bullen, M. (1998). Participation and critical thinking in online university distance education. *Journal of Distance Education, 13*(2), 1-32.

Butler, B., Sproull, L., Kiesler, S., & Kraut, R. (2002). Community effort in online groups: Who does the work and why. In S. Weisband (Ed.), *Leadership at a distance* (pp. 171-194). Hillsdale, NJ: Lawrence Erlbaum Associates.

Byrne, P. S., & Long, B. E. L. (1976). *Doctors talking to patients: A study of the verbal behaviour of general practitioners consulting in their surgeries*. United Kingdom, London: Her Majesty's Stationery Office.

Cabanero-Johnson, P. S., & Berge, Z. (2009). Digital natives: Back to the future of microworlds in a corporate learning organization. *The Learning Organization 16*(4), 290-297.

Chan, C. V., Matthews, L. A., & Kaufman, D. R. (2009). A taxonomy characterizing complexity of consumer eHealth literacy. In *AMIA annual symposium proceedings* (pp. 86-90). Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2815448

Chan, C. C., Tsui, M. S., Chan, M. Y., & Hong, J. H. (2002). Applying the structure of the observed learning outcomes (SOLO) taxonomy on student's learning outcomes: An empirical study. *Assessment & Evaluation in Higher Education, 27*(6), 511-527.

Chin, W. W. (1998). Issues and opinion on structural equation modeling. *MIS Quarterly, 22*(1), 7–16.

Chiu, C. M., Hsu, M. H., & Wang, E. T. G. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems, 42*(3), 1872-1888.

Chua, A. Y. K., & Banerjee, S. (2013). Customer knowledge management via social media: The case of Starbucks. *Journal of Knowledge Management, 17*(2), 237-249.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

Connor, E. (2009). Taking the pulse of health information seeking. *Journal of Electronic Resources in Medical Libraries, 6*(3), 230-235.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.

Cox, A. (2005). What are communities of practice? A comparative review of four seminal works. *Journal of Information Science, 31*(6), 527-540.

Creswell, J. (1998). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Dalkir, K. (2011). *Knowledge management in theory and practice* (2nd ed.). Cambridge, MA: MIT Press.

Davidson, C., & Voss, P. (2002). *Knowledge management: An introduction to creating competitive advantage from intellectual capital*. Auckland, New Zealand: Tandem Press.

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education, 46*(1), 6-28.

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2009). Structuring asynchronous discussion groups: The impact of role assignment and self assessment on students' levels of knowledge construction through social negotiation. *Journal of Computer Assisted Learning, 25*(2), 177-188.

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2010). Roles as a structuring tool in online discussion groups: The differential impact of different roles on social knowledge construction. *Computers in Human Behavior, 26*(4), 516-523.

DeAndrea, D. C., & Anthony, J. C. (2013). Online peer support for mental health problems in the United States: 2004-2010. *Psychological Medicine*, 1-12. doi:10.1017/S0033291713000172

Dennen, V. P. (2008). Looking for evidence of learning: Assessment and analysis methods for online discourse. *Computers in Human Behavior, 24*(2), 205-219.

Douven, I. (2011). Abduction. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford, CA: Stanford University. Retrieved from http://plato.stanford.edu/entries/abduction/#ExpAbd

Duggan, P. S., Geller, G., Cooper, L. A., & Beach, M. C. (2006). The moral nature of patient-centeredness: Is it "just the right thing to do"? *Patient Education and Counseling, 62*(2), 271-276.

Editors of the American Heritage Dictionaries (Ed.). (2007). *The American Heritage medical dictionary*. Boston, MA: Houghton Mifflin Harcourt.

Esquivel, A., Meric-Bernstam, F., & Bernstam, E. V. (2006). Accuracy and self correction of information received from an internet breast cancer list: Content analysis. *British Medical Journal, 332*(7547), 939-942.

Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., & Stern, A. (2004). Health related virtual communities and electronic support groups: Systematic review of the effects of online peer to peer interactions. *British Medical Journal, 328*(7449), 1166-1171.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299.

Fahy, P. J., Crawford, G., Ally, M., Cookson, P., Keller, V., & Prosser, F. (2000). The development and testing of a tool for analysis of computer-mediated conferencing transcripts. *Alberta Journal of Educational Research, 46*(1), 85-88.

Fisher, C., Lauría, E., Chengalur-Smith, S., & Wang, R. (2012). *Introduction to information quality*. Bloomington, IN: AuthorHouse.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Fox, S., & Jones, S. (2009). *The social life of health information* [Research report of the Pew Internet & American Life Project]. Retrieved from the Pew Research Center website: http://www.pewinternet.org/Reports/2009/8-The-Social-Life-of-Health-Information.aspx

Fraidaki, K., Pramatari, K., & Doukidis, G. (2012). Seeking information in social media: The case of the healthcare sector. In H. Rahman, A. Mesquita, I. Ramos, & B. Pernici (Eds.), *Knowledge and technologies in innovative information systems* (pp. 241-253). Berlin, Germany: Springer.

Furst, E. J. (1981). Bloom's taxonomy of educational objectives for the cognitive domain: Philosophical and educational issues. *Review of Educational Research, 51*(4), 441-453.

Garrison, D. R. (1991). Critical thinking and adult education: A conceptual model for developing critical thinking in adult learners. *International Journal of Lifelong Education, 10*(4), 287-303.

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2), 87-105.

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education 15*(1), 7-23.

Gauld, R. (2011). Factors associated with e-mail and internet use for health information and communications among Australians and New Zealanders. *Social Science Computer Review, 29*(1), 161-171.

Gebauer, J., Tang, Y., & Baimai, C. (2008). User requirements of mobile technology: results from a content analysis of user reviews. *Information Systems and e-Business Management, 6*(4), 361-384.

General Medical Council. (2013). *Doctors' use of social media* [Good medical practice guidelines]. Retrieved from the UK General Medical Council website: http://www.gmc-uk.org/Doctors__use_of_social_media.pdf_51448306.pdf

Goldberg, C. G., Berman, L., & Gusberg, R. J. (2010). Limitations of online information on abdominal aortic aneurysm. *International Journal of Vascular Medicine, 2010*, 1-6. doi:10.1155/2010/789198

Goodhue, D., Lewis, W., & Thompson, R. (2006). PLS, small sample size, and statistical power in MIS research. In R. Sprague (Ed.), *Proceedings of the 39th Annual Hawaii International Conference on on System Sciences* (pp. 1-10). Los Alamitos, CA: IEEE Computer Society Press.

Gorman, D., Horsburgh , M., & Abbott, M. (2009). *A review of how the training of the New Zealand health workforce is planned and funded: A proposal for a reconfiguration of the Clinical Training Agency* [Report of the Minister of Health's taskforce]. Retrieved from the Health Workforce New Zealand website: http://www.healthworkforce.govt.nz/sites/all/files/cta-review-aug09_0.pdf

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., Kyriakidou, O., & Peacock, R. (2005). Storylines of research in diffusion of innovation: A meta-narrative approach to systematic review. *Social Science & Medicine, 61*(2), 417-430.

Grol, R., De Maeseneer, J., Whitfield, M., & Mokkink, H. (1990). Disease-centred versus patient-centred attitudes: Comparison of general practitioners in Belgium, Britain and The Netherlands. *Family Practice, 7*(2), 100-103.

Guidry, J. J., Fagan, P., & Walker, V. (1998). Cultural sensitivity and readability of breast and prostate printed cancer education materials targeting African Americans. *Journal of the National Medical Association, 90*(3), 165-169.

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17*(4), 397-431.

Gustafson, D. H., Hawkins, R., Boberg, E., Pingree, S., Serlin, R. E., Graziano, F., & Chan, C. L. (1999). Impact of a patient-centered, computer-based health information/support system. *American Journal of Preventive Medicine, 16*(1), 1-9.

Haigh, C., & Costa, C. (2013). Information about assisted dying: An evaluation of web-based information resources. *Journal of Research in Nursing*. Advance online publication. doi:10.1177/1744987112471139.

Hair J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice Hall.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. The *Journal of Marketing Theory and Practice, 19*(2), 139-152.

Hara, N., Bonk, C. J., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional Science, 28*(2), 115-152.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32*(4), 1008-1015.

Helitzer, D., Hollis, C., Cotner, J., & Oestreicher, N. (2009). Health literacy demands of written health information materials: An assessment of cervical cancer prevention materials. *Cancer Control, 16*(1), 70-78.

Henri, F. (1992). Computer conferencing and content analysis. In C. O'Malley (Ed.), *Collaborative learning through computer conferencing* (pp. 117-136). Berlin, Germany: Springer.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing, 20*, 277-319.

Higgins, J. P., & Green, S. (Eds.). (2008). *Cochrane handbook for systematic reviews of interventions*. Chichester, United Kingdom: Wiley-Blackwell.

Hmelo-Silver, C. E. (2003). Analyzing collaborative knowledge construction: Multiple methods for integrated understanding. *Computers & Education, 41*(4), 397-420.

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

Hwang, A. S. (2003). Training strategies in the management of knowledge. *Journal of Knowledge Management, 7*(3), 92-104.

Jadad, A. R., & Gagliardi, A. (1998). Rating health information on the Internet. *JAMA: the Journal of the American Medical Association, 279*(8), 611-614.

Järvelä, S., & Häkkinen, P. (2002). Web-based cases in teaching and learning—the quality of discussions and a stage of perspective taking in asynchronous communication. *Interactive Learning Environments, 10*(1), 1-22.

Jena, A. B., & Goldman, D. P. (2011). Growing internet use may help explain the rise in prescription drug abuse in the United States. *Health Affairs, 30*(6), 1192-1199.

Johnson, C. M. (2001). A survey of current research on online communities of practice. *The Internet and Higher Education, 4*(1), 45-60.

Jordan, V. M. B., Jeffery, M., & Farquhar, C. M. (2006). The Cochrane Library is now freely available to all New Zealanders. *The New Zealand Medical Journal, 119*(1236). Retrieved from https://researchspace.auckland.ac.nz/handle/2292/9613

Kanuka, H., & Anderson, T. (2007). Online social interchange, discord, and knowledge construction. *The Journal of Distance Education, 13*(1), 57-74.

Katz, J. (1998, December 29). Luring the lurkers. *Slashdot*. Retrieved from http://news.slashdot.org/story/98/12/28/1745252/luring-the-lurkers

Kautz, K., & Thaysen, K. (2001). Knowledge, learning and IT support in a small software company. *Journal of Knowledge Management, 5*(4), 349-357.

Kent, M. L., & Taylor, M. (1998), Building dialogic relationships through the World Wide Web. *Public Relations Review, 24* (3), 321–334.

Khan, K. S., Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine, 96*(3), 118-121.

Kickbusch, I., Wait, S., & Maag, D. (2005). *Navigating health: The role of health literacy*. Retrieved from the International Longevity Centre—UK website: http://www.ilcuk.org.uk/index.php/publications/publication_details/navigating_health_the_role_of_health_literacy

Kickbush, I., & Maag, D. (2008). Health literacy. In K. Heggenhougen & S. Quah (Eds.), *International Encyclopedia of Public Health* (Vol. 3, pp. 204-211). London, United Kingdom: Elsevier.

Kim, J., & King, J. (2004). Managing knowledge work: Specialization and collaboration of engineering problem-solving. *Journal of Knowledge Management, 8*(2), 53-63.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Knight, S.-A., & Burn, J. (2005). Developing a framework for assessing information quality on the World Wide Web. *Informing Science: International Journal of an Emerging Transdiscipline, 8*(5), 159-172.

Kreuter, M. W., Lukwago, S. N., Bucholtz, D. C., Clark, E. M., & Sanders-Thompson, V. (2003). Achieving cultural appropriateness in health promotion programs: Targeted and tailored approaches. *Health Education and Behavior, 30*(2), 133-146.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage Publications.

Kucuk, M. (2010). Lurking in online asynchronous discussion. *Procedia—Social and Behavioral Sciences, 2*(2), 2260-2263.

Kvale, S. (1995). The social construction of validity. *Qualitative Inquiry*, 1(1), 19-40.

Laine, C., & Davidoff, F. (1996). Patient-centered medicine. *JAMA: The Journal of the American Medical Association, 275*(2), 152-156.

Lasker, J. N., Sogolow, E. D., & Sharim, R. R. (2005). The role of an online community for people with a rare disease: content analysis of messages posted on a primary biliary cirrhosis mailinglist. *Journal of Medical Internet Research, 7*(1), e10. doi:10.2196/jmir.7.1.e10

Lawshe, C. H. (1975). A Quantitative approach to content validity. *Personnel Psychology, 28*(4), 563-575.

Leahey, T. H., & Harris, R. J. (1985). *Human learning*. Englewood Cliffs, NJ: Prentice-Hall.

Lee, C., Gray, S. W., & Lewis, N. (2010). Internet use leads cancer patients to be active health care consumers. *Patient Education and Counseling, 81*(Supplement 1), S63-S69.

Lee, J. Y., & Sundar, S. S. (2012). To tweet or to retweet? That is the question for health professionals on twitter. *Health Communication*, 1-16. doi:10.1080/10410236.2012.700391

Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information & Management, 40*(2), 133-146.

Lefrançois, G. R. (1995). *Theories of human learning* (3rd ed.). Pacific Grove, CA: Brooks.

Lin, J., Lu, Y., Wang, B., & Wei, K. K. (2011). The role of inter-channel trust transfer in establishing mobile commerce trust. *Electronic Commerce Research and Applications, 10*(6), 615-625.

Lipkin, M., Quill, T. E., & Napodano, R. J. (1984). The medical interview: A core curriculum for residencies in internal medicine. *Annals of Internal Medicine, 100*(2), 277-284.

Magnier-Watanabe, R., Yoshida, M., & Watanabe, T. (2010). Social network productivity in the use of SNS. *Journal of Knowledge Management, 14*(6), 910-927.

Malhotra, N. K., Kim, S. S., & Patil, A. (2006). Common method variance in IS research: A comparison of alternative approaches and a reanalysis of past research. *Management Science, 52*(12), 1865-1883.

Maloney-Krichmar, D., & Preece, J. (2005). A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Transactions on Computer-Human Interaction, 12*(2), 201-232.

Mancuso, J. M. (2008). Health literacy: A concept/dimensional analysis. *Nursing & Health Sciences, 10*(3), 248-255.

Manganello, J. A. (2008). Health literacy and adolescents: A framework and agenda for future research. *Health Education Research, 23*(5), 840-847.

Mankoff, J., Kuksenok, K., Kiesler, S., Rode, J. A., & Waldman, K. (2011). *Competing online viewpoints and models of chronic illness*. In Proceedings of the 2011 annual conference on Human factors in computing systems (pp. 589-598). New York, NY: Association for Computing Machinery.

Mansfield, S. J., Morrison, S. G., Stephens, H. O., Bonning, M. A., Wang, S.-H., Withers, A. H. J., & Perry, A. W. (2011). Social media and the medical profession. *The Medical Journal of Australia, 194*(12), 642-644.

Marcoulides, G. A., & Saunders, C. (2006). PLS: a silver bullet? *MIS Quarterly, 30*(2), 3-9.

Marshall, C., & Rossman, G. B. (2010). *Designing qualitative research*. Thousand Oaks, CA: Sage.

McAdam, R., & McCreedy, S. (1999). The process of knowledge management within organizations: A critical assessment of both theory and practice. *Knowledge and Process Management, 6*(2), 101-113.

McAdam, R., & Reid, R. (2001). SME and large organisation perceptions of knowledge management: Comparisons and contrasts. *Journal of Knowledge Management, 5*(3), 231-241.

McCray, A. T. (2005). Promoting health literacy. *Journal of the American Medical Informatics Association, 12*(2), 152-163.

McKinney, V., Yoon, K., & Zahedi, F. M. (2002). The measurement of web-customer satisfaction: An expectation and disconfirmation approach. *Information Systems Research, 13*(3), 296.

McWhinney, I. R. (1985). Patient-centred and doctor-centred models of clinical decision making. In M. Sheldon, J. Brook, & A. Rector (Eds.), *Decision making in general practice* (pp. 31-46). London, United Kingdom: Stockton.

Mead, N., & Bower, P. (2000). Patient-centredness: A conceptual framework and review of the empirical literature. *Social Science & Medicine 51*(7), 1087–1110.

Mehta, Nikhil. (2008). Successful knowledge management implementation in global software companies. *Journal of Knowledge Management, 12*(2), 42-56.

Milne, P., & Callahan, S. (2006). ActKM: The story of a community. *Journal of Knowledge Management, 10*(1), 108-118.

Moore, J. L., & Marra, R. M. (2005). A comparative analysis of online discussion participation protocols. *Journal of Research on Technology in Education, 38*(2), 22.

Muthen, L. K., & Muthen, B. O. (2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

Myers, M. D. (1997). Qualitative research in information systems. *MIS Quarterly, 21*(2), 241-242.

Nagel, L., Blignaut, A. S., & Cronje, J. C. (2009). Read-only participants: A case for student communication in online classes. *Interactive Learning Environments, 17*(1), 37-51.

Newman, D. R., Johnson, C., Webb, B., & Cochrane, C. (1997). Evaluating the quality of learning in computer supported co-operative learning. *Journal of the American Society for Information Science, 48*(6), 484-495.

Nicolaou, A. I., & McKnight, D. H. (2006). Perceived information quality in data exchanges: Effects on risk, trust, and intention to use. *Information Systems Research, 17*(4), 332-351.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. New York, NY: Oxford University Press.

Nonnecke, B., & Preece, J. (2000). Lurker demographics: Counting the silent. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 73-80). New York, NY: Association for Computing Machinery.

Nordfeldt, S., Ãngarne-Lindberg, T., & Bertero, C. (2012). To use or not to use—practitioners' perceptions of an open web portal for young patients with diabetes. *Journal of Medical Internet research, 14*(6), e154. doi:10.2196/jmir.1987

Norman, C. D., Jackson, S., Larkin, J., Lombardo, C., Meyer, T., & Skinner, H. (2008). *The five E's for health: Understanding eLiteracies for eLearning in electronic educational environments* (Research report for Canadian Council on Learning). Retrieved from the Canadian Council on Learning website: http://www.ccl-cca.ca/ccl/Research/FundedResearch/200812FiveEsforHealth.html

National Health IT Board. (2010). *National health IT plan: Enabling an integrated healthcare model* [Discussion document]. Retrieved from the New Zealand National Health IT Board website: http://www.ithealthboard.health.nz/sites/all/files/National Health IT Plan v11_1.pdf

November, N., & Day, K. (2012). Using undergraduate's digital literacy skills to improve their discipline-specific writing: A dialogue. *International Journal for the Scholarship of Teaching and Learning, 6*(2), 1-21. Retrieved from http://academics.georgiasouthern.edu/ijsotl/v6n2.html

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Nutbeam, D. (1998). Health promotion glossary. *Health promotion international, 13*(4), 349-364.

Nutbeam, D. (2000). Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century. *Health promotion international, 15*(3), 259-267.

Orlikowski, W. J., & Robey, D. (1991). Information technology and the structuring of organizations. *Information Systems Research, 2*(2), 143-169.

Paasche-Orlow, M. K., & Wolf, M. S. (2007). The causal pathways linking health literacy to health outcomes. *American Journal of Health Behavior, 31*(Supplement 1), S19-S26.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Reviews in Psychology, 49*(1), 345-375.

Park, H., Rodgers, S., & Stemmle, J. (2013). Analyzing health organizations' use of Twitter for promoting health literacy. *Journal of Health Communication, 18*(4), 410-425.

Paroutis, S., & Al Saleh, A. (2009). Determinants of knowledge sharing using Web 2.0 technologies. *Journal of Knowledge Management, 13*(4), 52-63.

Pena-Shaff, J. B., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education, 42*(3), 243-265.

Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational researcher, 24*(7), 5-12.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903.

Porter, A., & Edirippulige, S. (2007). Parents of deaf children seeking hearing loss-related information on the internet: The Australian experience. *Journal of Deaf Studies and Deaf Education, 12*(4), 518-529.

Post, R. E, & Mainous, A. G. (2010). The accuracy of nutrition information on the Internet for type 2 diabetes. *Archives of Internal Medicine, 170*(16), 1504-1506.

Powell, J., Inglis, N., Ronnie, J., & Large, S. (2011). The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. *Journal of Medical Internet research, 13*(1), e20. doi:10.2196/jmir.1600

Preece, J., & Maloney-Krichmar, D. (2003). Online communities: Focusing on sociability and usability. In J. A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook* (pp. 596-620). Mahwah, NJ: Lawrence Erlbaum.

Puntambekar, S. (2006). Analyzing collaborative interactions: Divergence, shared understanding and construction of knowledge. *Computers & Education, 47*(3), 332-351.

Rajagopalan, M. S., Khanna, V. K., Leiter, Y., Stott, M., Showalter, T. N., Dicker, A. P., & Lawrence, Y. R. (2011). Patient-oriented cancer information on the internet: A comparison of Wikipedia and a professionally maintained database. *Journal of Oncology Practice, 7*(5), 319-323.

Relevant. (n.d.). In *Merriam-Webster Online*. Retrieved from http://www.merriam-webster.com/dictionary/relevant

Rienties, B., Giesbers, B., Tempelaar, D., Lygo-Baker, S., Segers, M., & Gijselaers, W. (2012). The role of scaffolding and motivation in CSCL. *Computers & Education 59*(3), 893-906.

Riffe, D., Lacy, S., Fico, F., & Fico, F. G. (2005). *Analyzing media messages: Using quantitative content analysis in research.* Mawah, NJ: Erlbaum.

Robb, G., & Seddon, M. (2006). Quality improvement in New Zealand healthcare. Part 6: Keeping the patient front and centre to improve healthcare quality. *Journal of the New Zealand Medical Association, 119*(1242), 1-9.

Rosenberg, W., & Donald, A. (1995). Evidence based medicine: An approach to clinical problem-solving. *British Medical Journal, 310*(6987), 1122-1126.

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development, 52*(1), 5-18.

Rowley, J. (2000). From learning organisation to knowledge entrepreneur. *Journal of Knowledge Management, 4*(1), 7-15.

Sale, J. E. M., Lohfeld, L. H., & Brazil, K. (2002). Revisiting the quantitative-qualitative debate: Implications for mixed-methods research. *Quality & Quantity, 36*(1), 43-53.

Salmon, G. (2003). *E-moderating: The key to teaching and learning online*. London, United Kingdom: Taylor & Francis.

Salomon, G. (1993). No distribution without individuals' cognition: A dynamic interactional view. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 111-138). New York, NY: Cambridge University Press.

Schellens, T., & Valcke, M. (2005). Collaborative learning in asynchronous discussion groups: What about the impact on cognitive processing? *Computers in Human Behavior, 21*(6), 957-975.

Schellens, T., & Valcke, M. (2006). Fostering knowledge construction in university students through asynchronous discussion groups. *Computers & Education, 46*(4), 349-370.

Schrire, S. (2004). Interaction and cognition in asynchronous computer conferencing. *Instructional Science, 32*(6), 475-502.

Secko, D. M., Tlalka, S., Dunlop, M., Kingdon, A., & Amend, E. (2011). The unfinished science story: Journalist–audience interactions from the Globe and Mail's online health and science sections. *Journalism, 12*(7), 814-831.

Senge, P. M. (2006). *The fifth discipline: The art & practice of the learning organization*. New York, NY: Doubleday.

Sloep, P., & Kester, L. (2009). From lurker to active participant. In R. Koper (Ed.), *Learning network services for professional development* (pp. 17-25). Berlin, Germany: Springer.

Sorensen, K., Van Den Broucke, S., Brand, H., Fullam, J., Doyle, G., Pelikan, J., & Slonszka, Z. (2012). Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health, 12*(80), 1-13. Retrieved from http://www.biomedcentral.com/1471-2458/12/80

Stahl, G. (2004). Building collaborative knowing: Elements of a social theory of CSCL. In J.-W. Strijbos, P. Kirschner, & R. Martens (Eds.), *What we know about CSCL* (pp. 53–86). Boston, MA: Kluwer.

Straub, D., Boudreau, M. C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information Systems, 13*(24), 380-427.

Suggs, L. S. (2006). A 10-year retrospective of research in new technologies for health communication. *Journal of Health Communication, 11*(1), 61-74.

Taber, K. S. (2000). Case studies and generalizability: Grounded theory and research in science education. *International Journal of Science Education, 22*(5), 469-487.

Takahashi, M., Fujimoto, M., & Yamasaki, N. (2003). The active lurker: Influence of an in-house online community on its outside environment. In *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work* (pp. 1-10). New York, NY: Association for Computing Machinery.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis, 48*(1), 159-205.

Veerman, A. L., Andriessen, J. E. B., & Kanselaar, G. (1999). Collaborative learning through computer-mediated argumentation. In C. Hoadley & J. Roschelle (Eds.), *Proceedings of the 1999 Conference on Computer Support for Collaborative Learning* (pp. 640 - 650). Palo Alto, CA: Stanford University.

Veerman, A. L., & Veldhuis-Diermanse, A. E. (2001). Collaborative learning through computer-mediated communication in academic education. In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.), *European perspectives on computer-supported collaborative learning* (pp. 625-632). Maastricht, Netherlands: University of Maastricht.

Vuori, V., & Okkonen, J. (2012). Knowledge sharing motivational factors of using an intra-organizational social media platform. *Journal of Knowledge Management, 16*(4), 592-603.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes* (14th ed.) [Collection of essays originally published in Russian between 1935 and 1966]. (A. R. Luria, M. Lopez-Morillas, & M. Cole, Trans.). Cambridge, MA: Harvard University Press.

Wang, R. Y, & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of management information systems 12*(4), 5-33.

Waters, R. D., Canfield, R. R., Foster, J. M., & Hardy, E. E. (2011). Applying the dialogic theory to social networking sites: Examining how university health centers convey health messages on Facebook. *Journal of Social Marketing, 1*(3), 211-227.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly, 26*(2), 13-23.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education, 46*(1), 71-95.

Wenger, E. (1998, June). Communities of practice: Learning as a social system. *Systems Thinker, 9*(5), 1-8.

Wetzels, M., Odekerken-Schroder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS Quarterly, 33*(1), 177-195.

White, M., & Dorman, S. M. (2001). Receiving social support online: Implications for health education. *Health Education Research, 16*(6), 693-707.

Whittaker, S., Isaacs, E., & O'Day, V. (1997). Widening the net: Workshop report on the theory and practice of physical and network communities. *ACM SIGCHI Bulletin, 29*(3), 27-30.

Winefield, H., Murrell, T., Clifford, J., & Farmer, E. (1996). The search for reliable and valid measures of patient-centredness. *Psychology and Health, 11*(6), 811-824.

Wright, K. B., & Bell, S. B. (2003). Health-related support groups on the Internet: Linking empirical findings to social support and computer-mediated communication theory. *Journal of Health Psychology, 8*(1), 39-54.

Wright, S. (2009). The role of the moderator: Problems and possibilities for government-run online discussion forums. In T. Davies & S. P. Gangadharan (Eds.), *Online deliberation: Design, research, and practice* (pp. 233-242). Stanford, CA: CSLI Publications.

Van Woerkum, C. M. (2003). The Internet and primary care physicians: Coping with different expectations. *The American Journal of Clinical Nutrition, 77*(4), 1016S-1018S.

Yin, R. K. (2009). *Case study research* (4th ed.). Los Angeles, CA: Sage.

Yousuf, M. I. (2007). Using experts' opinions through Delphi technique. *Practical Assessment, Research & Evaluation, 12*(4), 1-8.

Zarcadoolas, C., Pleasant, A., & Greer, D. S. (2005). Understanding health literacy: An expanded model. *Health Promotion International, 20*(2), 195-203.