

## **4 Research Methodology**

Chapter 2 was dedicated to a review of the knowledge management literature, to providing a theoretical background for this thesis, and setting out the research questions. Chapter 3 focused on the issues of knowledge management systems and the relevant analytical frameworks, and developed research models and a set of research hypotheses. This chapter addresses the critical methodological issues, including the philosophies, approaches, and methods, to be used in this thesis. It is organized as follows. Section 4.2 addresses the IS research approaches and explains why the positivist survey approach is chosen as being appropriate for this research; Section 4.3 discusses the research methods appropriate to this study; Section 4.4 is dedicated to survey instrument development; Section 4.5 presents a preliminary study of the survey instruments, and the development of the final questionnaire. Section 4.6 sets out how the survey was conducted, and Section 4.7 summarizes the Chapter.

### **4.1 Introduction**

The main purpose of this research is to empirically investigate why people within organizations accept and use KMS, and the impact of knowledge management systems on organizational knowledge management soft infrastructure (e.g., social capital). As KMS are generally considered to be an application of information and communication technologies, it was felt useful to review typical research methodologies used in previous studies of information systems, to assist in the selection of an appropriate methodology and research design for this study.

### **4.2 Research Approaches and Assumptions in IS Research**

Information systems research usually involves studying the interaction of information technologies and humans in organizational contexts at both macro and micro levels of analysis (Orlikowski and Baroudi 1991). A variety of research approaches can be found in the information systems research literature, including:

- Field experiments and laboratory experiments (Benbasat 1989).
- Surveys (Kraemer 1991; Newsted, Huff et al. 1998).
- Case studies (Cash and Lawrence 1989; Lee 1989; Markus and Lee 1999).

- Action research (Baskerville and Wood-Harper 1998).
- Longitudinal research (Venkatesh and Davis 2000).
- Historical research (Mason, McKeen et al. 1997; Mason, McKeen et al. 1997).
- Protocol analysis and ethnography (Orlikowski and Baroudi 1991; Schultze 2000).

According to Orlikowski and Baroudi (1991), research enquires involve three philosophical assumptions, namely:

- The underlying nature of the phenomena under investigation.
- The use of appropriate research method, and
- The nature of valid evidence.

Three popular research philosophies found in IS include:

- Positivist (Westland 2004).
- Interpretive (Myers 1997; Markus and Lee 1999).
- Critical (Taylor and Trujillo 2001).

Orlikowski and Baroudi (1991) proposes that these research perspectives can be analysed using a framework of three sets of beliefs, that is,

- Beliefs about the phenomenon or object of inquiry, including an objective or subjective view of physical and social reality, the assumptions of human rationality, and assumptions concerning social relations and interactions.
- Beliefs about knowledge, that is the criteria by which valid knowledge about a phenomenon may be constructed and evaluated, and which research methods and techniques are appropriate for the collection of valid empirical evidence, and
- Beliefs about the relationship between theory and practice, that is, the role of theory in practice, and the values and intentions that researchers bring to their work.

The following sub-sections explore briefly the three research philosophies that have been widely used in information systems research.

### **4.2.1 The Positivist Philosophy of Information Systems Research**

With roots in logical positivism and natural science research tradition, the positivist research perspective is based on the existence of a priori fixed relationships within phenomena. The positivist philosophy reflects a set of precepts which assume that the phenomenon of interest exists independent of observers or researchers, and its nature can be understood, characterized, and measured relatively un-problematically.

Any chosen aspects of the phenomenon of inquiry can always be relatively rigorously represented by the researcher's model, with an appropriate set of constructs which are precise and which correspond one-to-one to the events, objects, or features in physical and social reality, and which have fixed and invariant meanings.

The unidirectional cause-effect relationships embedded in the researcher's model are capable of being tested via hypothetic-deductive logic and analysis, and finally the study can be value-free.

From the positivist perspective, human action is considered intentional and rational, or at the least, boundedly rational. The social relations and human interactions are assumed to be relatively stable and orderly. With respect to knowledge, the positivist perspective is concerned with the empirical testability of theories. For example, a theory is considered true and acceptable only if it is repeatedly not falsified by empirical events (Orlikowski and Baroudi 1991). The positivist perspective suggests that research is conducted via hypothetic-deductive logic and analysis, which is the basis of the generated knowledge. Once universal laws or principles concerning certain phenomenon are established, relevant low-level hypotheses may be deduced from them, which then enable explanation, prediction and control of certain events and actions across situations. In order to conduct valid positivist research, it is held to be necessary to follow one of a given set of research methodologies. Among these methodologies, sample surveys and controlled experiments are considered as the primary data collection techniques, and inferential statistics is the data analysis method used to test and verify causal laws. For a positivist research practice, the following are considered crucial:

- The model development methodology (Glymour and Spirtes 1988; Lee, Barua et al. 1997),

- Construct development (i.e., refinement and operationalization)(Bagozzi, Yi et al. 1991; Schriesheim, Powers et al. 1993; Hox 1997; Bagozzi and Edwards 1998),
- The validity and reliability of instruments (Churchill 1979; Straub 1989; Straub, Boudreau et al. 2004),
- Researcher detachment from the research process (Orlikowski and Baroudi 1991),
- The random assignment of subjects (Orlikowski and Baroudi 1991),
- Control over confounding influences (Orlikowski and Baroudi 1991),.

The positivist perspective suggests that the relationship between theory and practice is essentially technical. As value-neutrality is taken for granted by positivists, this means that researchers can comment on means, but not ends. In addition, the assumption of value-free research also suggests a detachment between researchers and the phenomena under inquiry.

In sum, as a dominant research perspective in information systems research, the positivist research perspective plays a significant role in information systems research (Kraemer 1991; Orlikowski and Baroudi 1991; Westland 2004). While the positivist research perspective, as a well-established scientific research approach, is well-accepted and in widespread use by information systems researchers (Palvia, Mao et al. 2003), it has limitations, and has been criticized for its intrinsic pitfalls (Orlikowski and Baroudi 1991; Lee 1999). For example (Orlikowski and Baroudi 1991; Lee 1999):

- The positivist research perspective implicates a mechanical and static worldview which deals with phenomenon of interest as stable, fragmentable, and invariant (Orlikowski and Baroudi 1991);
- The application of positivist precepts to the study of real social phenomena may be problematic, because these precepts are ideals, over-simplified versions of the complex physical and social reality ( Mingers 2001);
- Not all of the social phenomena of interest can be represented by causal-effect models and unidirectional causal-effect relationships; some social phenomena are so complex and dependent upon time, contexts, and people that the positivist approach may not be applicable to them ( Miller 2001).

### **4.2.2 The Interpretive Philosophy of Information Systems Research**

The interpretive perspective is based on social constructionism, which asserts that reality and knowledge are social constructed and are thus incapable of being understood independently of the social actors that construct and make sense of that reality (Orlikowski and Baroudi 1991; Myers 1997; Creswell 1998; Taylor and Trujillo 2001). From the interpretive perspective, the world is conceived of as an emergent social process, an extension of human consciousness and subjective experience, instead of as a fixed set of objects. The aim of interpretive research is to

- Interpret how the members of a social group, through their participation in social processes, enact their particular realities and give meaning to them, and
- Demonstrate how these meanings, beliefs and intentions help to constitute the members' social actions (Orlikowski and Baroudi 1991).

The interpretive research perspective believes that social realities such as organizations, groups, social networks and relations and technology systems do not exist independently of people, and therefore can only be interpreted, not apprehended, characterized, and measured in some objective or universal way (Orlikowski and Baroudi 1991; Myers 1997). From the interpretive perspective, social reality is understood to be constructed and reconstructed through ongoing social interactions and actions (Orlikowski and Baroudi 1991). Subjective meanings, social-political and symbolic actions are considered as the most important factors in understanding the social interactions and actions. Through their socialization into, interaction with, and participation in a social world, people construct their realities and give them status and meaning (Creswell 1998). The interpretive research perspective attempts to reveal how and why these social realities are formed and reformed (Orlikowski and Baroudi 1991; Orlikowski 2000). In addition, as meanings are the result of negotiation, the interpretation of reality may shift over time as circumstances, objectives, and constituencies change (Orlikowski and Baroudi 1991).

With regard to the epistemological beliefs related to interpretive research, the interpretive perspective believes that the social process is not captured in prior researcher-devised constructs, simplified and objective causal-effect models (i.e., uni-directional and linear causal-effect relationships), hypothetical deductions, covariance,

and degrees of freedom. Instead, interpretive study posits reciprocally interacting models of causality, with the intentions of interpreting actors' views and meanings of the phenomenon of interest and their role in it, so as to construct explanations accounting for the way that subjective meanings are created and sustained in a particular setting. In addition, the language that humans use to enact and describe social practices is considered to be an inseparable part of these social practices (Orlikowski and Baroudi 1991; Creswell 1998). Thus,

*"Understanding social reality requires understanding how practices and meanings are formed and informed by the language and tacit norms shared by humans working towards some shared goal" (Orlikowski and Baroudi 1991, p. 14).*

Field studies are considered to be appropriate research methods for generating valid interpretive knowledge (Orlikowski and Baroudi 1991; Klein and Myers 1999). Field studies examine the phenomena of interest and humans within their social settings, allow participants to use their own language, images, concepts and experiences, and attempt to describe, interpret, analyze, and understand the social world from the participants' perspective (Snow and Thomas 1994; Orlikowski 2000). Interpretive research attempts to derive valid constructs and categories from the field, instead of imposing externally defined categories and constructs on the phenomenon of interest (Creswell 1998; Taylor and Trujillo 2001).

With regard to the relationship between theory and practice, interpretive research can never assume a value-neutral stance; the study is always shaped by the researchers' prior assumptions, beliefs, values, and interests (Orlikowski and Baroudi 1991). The interpretive perspective can be differentiated by the role of the researcher in investigating phenomena, for example, a "weak" interpretive approach suggests that the researcher is presumed to describe a phenomenon in words and categories of actors through various data collection techniques, while a "strong" interpretive approach suggests that the researcher attempts to understand the phenomenon of interest by both enacting the social reality being studied and describing the phenomenon with words and categories from actors (Orlikowski and Baroudi 1991). The "weak" interpretive perspective can be used as a complement to positivist research, for instance, by

generating hypotheses for further investigation, whilst the role of a “strong” interpretive approach is to replace the positivist approach, as this approach is considered to be based on fundamentally incompatible philosophical assumptions (Orlikowski and Baroudi 1991).

In sum, the interpretive research perspective is explicitly designed to capture complex, dynamic, context- and time-dependent social phenomena. Interpretive research focuses on investigating how humans enact a shared social reality and come to share a set of meanings around this reality, through understanding human behaviour from the perspective and intentions of the human actors themselves (Orlikowski and Baroudi 1991). Field study is the appropriate method to conduct successful interpretive research. However, the interpretive research approach is subject to criticism for its intrinsic pitfalls and limitations. For example, the interpretive perspective:

- Does not examine the external conditions which may give rise to certain meanings and experiences.
- Omits to explain the unintended consequences of human action which are often significant forces in shaping social reality.
- Ignores the contradictions which may be endemic to social systems, and
- Neglects to explain the historical changes to social systems (Orlikowski and Baroudi 1991).

#### **4.2.3 The Critical Philosophy of Information Systems Research**

As a distinctive research perspective, the critical researcher seeks to evaluate and transform the social reality under investigation (Orlikowski and Baroudi 1991; Creswell 1998; Taylor and Trujillo 2001). The critical perspective focuses on critiquing the existing social realities and on revealing any contradictions and conflicts that may exist within their structures, for the purposes of evaluation and transformation. It does not attempt to explain or predict the status quo.

The critical perspective is based on the belief that social reality is historically constituted, and that social reality can only be understood by examining its development process. Although the critical perspective believes that people have the potential to

change their material and social circumstances, it recognizes that people's capacity to enact change is constrained by the prevailing systems of economic, political, and cultural authority (Orlikowski and Baroudi 1991). Therefore, the critical research approach attempts to identify and understand the various forms of social domination, so that people can act to change or remove them. The critical perspective emphasizes a dynamic and systematic view of phenomena of interest, that is, that things can never be treated as isolated elements. The dialectical relationship between elements and their totality (elements and totality are inseparable and bound by an essential interdependence), is an important focus in the critical perspective; it seeks to understand how this relationship is shaped by historical and contextual conditions (Creswell 1998). The critical perspective assumes that a new social reality emerges from the inequalities and conflicts caused by the inherent and fundamental contradictions in existing social relationships and practices within societies and organizations. As these contradictions are the result of the incompatible development of the parts which constitute the totality, and can never be avoided, the critical perspective pays particular attention to the development process of phenomena (Orlikowski and Baroudi 1991).

With regard to knowledge, the critical perspective believes that knowledge is grounded in social and historical practices (Orlikowski and Baroudi 1991). Therefore, critical studies tend to be longitudinal, and long-term historical studies (Mason, McKeen et al. 1997; Mason, McKeen et al. 1997) and ethnographic studies are considered as appropriate research methods for critical research (Schultze 2000). Critical researchers do not only accept the self-understanding of participants, but critically analyze it through the theoretical framework which they adopt to conduct their research (Orlikowski and Baroudi 1991).

The relationship between theory and practice in critical research lies in the role of the researcher in revealing the contradictions underlying the social reality being studied, thereby initiating change in the social relations and practices, and helping to remove the bases of alienation and domination. Attempting to change social relations and practices is considered to be the aim of critical research (Orlikowski and Baroudi 1991). For example, the aim of critical organization research is to develop an organization science capability of changing organizational processes (Steffy and Grimes 1986)



While the critical research approach offers an insightful perspective on the phenomena of interest in information systems research, it suffers from some weaknesses (Creswell 1998). As the critical research perspective focuses on the assumption that contradiction is a function of socio-economic class societies, and on initiating social change, it may thus ignore the development of theory. As argued by Orlikowski and Baroudi (1991), critical researchers often are not critical enough of their own concepts and theoretical models. In addition, whether the lack of common philosophical standards for the evaluation of theories leads to the production of acceptable theories is still debatable (Orlikowski and Baroudi 1991).

#### **4.2.4 Summary**

In this section, three of the main research philosophies in management information systems research, i.e., the positivist perspective, the interpretive perspective, and the critical perspective have been reviewed. Each research philosophy has been examined in terms of its beliefs about the nature of the phenomenon of interest, knowledge standards and appropriate research methods, and the relationship between theory and practice. This discussion serves as a framework for choosing the suitable and viable research approach(s) for the phenomenon of interest. The following sections will discuss the main research methods and techniques chosen for the research undertaken for this thesis.

### **4.3 Research Methods for this Study**

As discussed in the previous section, the three research approaches involve three sets of different research methods and techniques. For instance, while survey and experimental methods are considered appropriate for positivist research (Benbasat 1989; Attewell and Rule 1991; Kraemer 1991), case study is believed to be the right way to conduct an interpretive research (Cash and Lawrence 1989; Lee 1989). Research methods and techniques are critical to the desired research outcomes because the generation of valid knowledge relies on the choice of appropriate methods and tools.

#### **4.3.1 The Positivist Approach and Survey Research**

This study is seeking to explore and establish and verify the relationships between the

use of KMS and its impacts on end-users. Positivist approach is considered to be appropriate methodology to achieve the target with a great rigor and very promising results.

In order to conduct valid positivist research, a number of research methodologies are considered to be appropriate, and following these methodologies is considered as the only way in which valid knowledge can be obtained. Among the methodologies, sample surveys are considered as the primary data collection technique (Kraemer 1991), and inferential statistics is the data analysis method used to test and verify causal laws (Tabachnick and Fidell 1996).

Survey research is one of the most common research methods used in information systems research (Kraemer 1991; Lucas 1991). For instance, Orlikowski and Baroudi (1991), in an examination of five years of information systems literature published between January 1983 and May 1988 in four major information systems journals, found that survey method was used in about 50% of papers. More recently, in an investigation on the research methodologies use by research articles in seven selected leading academic MIS journals between 1993 and 2003 (Palvia, Mao et al. 2003, 2004) found that surveys were the most used methodology.

Survey research has many advantages. It is especially well suited to cases in which the generalizing of results from a sample to a population is important (Gutek 1991), and an appropriate choice of method for the objective verification of hypotheses. Survey research can accurately document the norm, identify extreme outcomes, and delineate associations between variables in a sample. As suggested by Attewell and Rule (1991), survey research usually serves as a methodology of verification rather than of discovery. In addition, survey studies are relatively unobtrusive compared to other research methods.

Survey research involves the systematic gathering of information for scientific purposes from more than a few entities using standardized instruments, and the performance of statistical analysis of the information (Kraemer and Dutton 1991; Lucas 1991). The aim

of survey research is to generalize from the sample to the population on some substantive issue.

According to the literature, survey research has three distinct characteristics.

- Firstly, it is designed to produce quantitative descriptions of certain aspects of a population being studied.
- Secondly, the principal means of gathering information is by asking structured questions.
- Thirdly, information is generally collected about only a fraction of the study population, i.e., a sample, and is collected in such a way so as to permit the generalization of findings to the population.

Survey research is primarily concerned either with relationships between variables or with the projection of findings descriptively to a predefined population. Survey research involves subjects, such as individuals, groups, organizations or communities, from whom standardized information is collected. The size of the sample has to be large enough to allow statistical analysis (Kraemer 1991; Hinkin 1998).

- Survey research requires
- An explicit theoretical framework or model,
- A research design,
- Data collection instruments related to the theoretical constructs in the framework.
- A sampling strategy that allow generalization of research findings (Lucas 1991; Sekaran 1992).

Lucas (1991) suggested a survey research process model which identified a set of important steps for conducting a quality survey research, namely

- Formulate a research question.
- Develop a research design.
- Conduct the analysis, and
- Write up the research.

The research question is at the centre of information systems survey research. According to Lucas (1991), the research question involves defining the research area, locating relevant theories, developing a research model, and delineating hypotheses. Prior to this, the researcher needs to identify a significant and interesting research problem (Porter 1991). Once this is found, the researcher has to find relevant theories, although these are not always available. As argued by Lucas (1991a, 1991b), failure to find a theory often leads to survey research which lacks hypotheses testing and rigor. However, the researcher can overcome the problem of lack of available theories by developing a research model based on existing studies (Lucas 1991; Lee, Barua et al. 1997). Development of a research model requires balancing the complexity of the model (e.g., the number of variables) and the practicality of collecting sufficient data to test it. Hypothesis development is another important issue in information systems survey research. A research model can play an important role in moving from theory to hypotheses. The hypotheses should be delineated from the researcher's research model or theory and be significant. In practice, a single hypothesis may generate several operational hypotheses for testing (Lucas 1991; Lee, Barua et al. 1997).

Research design is considered to be a crucial determinant of the credibility of a study (Kraemer and Dutton 1991; Lucas 1991; Sekaran 1992). The research design includes determining how to measure variables in the research model, that is

- How to operationalize the variables,
- How to design or choose measurement instruments, and
- How to confirm the reliability and validity of the measurement instruments

It also includes identifying the population to sample and the unit of analysis, and the method of data collection. For a survey research where generalizability is important, a survey is only as good as its sample (Gutek 1991). While random samples are considered desirable, however, a researcher often has to use entirely opportunistic samples (or "convenience" samples) or purposive samples due to the limitations of available resources, such as accessibility to the target organizations (King 1991). On the other hand, for many studies, their principal goal is to test the strength of the relationships among a set of variables of interests, instead of generalizing their findings to the target population. For such studies, opportunistic samples (or "convenience"

samples) or purposive samples may be used (Gutek 1991). In addition, as a great deal of data collection in information systems survey research is cross-sectional, King (1991) argued that cross-sectional survey research can also provide causal inference, provided the work is done correctly (i.e., an appropriate logic of analysis and inquiry). Furthermore, the level of the sampling unit may range from individual people to work groups to firms (organizations) to industries to nation states depending upon the research question (Gutek 1991).

There are three common methods of data collection,

- Questionnaires filled out by respondents.
- Face-to-face interviews in which the interviewer completes the instrument, and
- Telephone interviews in which the interviewer completes the instrument.

Each of them has its own advantages and drawbacks, and the methods can be used alone or together (Zikmund 2000).

Once the data collection has been completed, data analysis will allow the researcher to test the research model and hypotheses. At this stage, the researcher must decide which statistics to use. A range of statistical analysis techniques are available for IS research. Although structural equation modelling (SEM) have become a popular analysis technique for MIS research (please refer to subsection 4.3.4. for details about SEM), the choice of a single or multiple techniques depends on several factors, including the research question and purpose, the nature of the research, and sample data constraints (e.g., sample size, sample distribution assumptions). For instance, covariance-based SEM is best suited for confirmatory study and theory testing, while partial least squares-based (PLS) SEM is better for predictive analysis and theory building (Gefen, Straub et al. 2000). However, linear regression can be a good choice in the case where critical SEM assumptions are violated, such as multi-collinearity (Tabachnick and Fidell 1996), and polynomial relationships (Gefen, Straub et al. 2000). The sample size also affects decisions on statistic analysis and hypothesis testing. The required minimal sample size for different inferential statistics may vary; for instance, for linear regression there should be at least 10 observations per independent variables in the equation, for co-variance-based structural equation modeling (e.g., LISREL), the

minimal sample size should be at least 100 – 150 cases, and for the partial least square-based structural equation modeling (PLS), at least 10 times the number of items in the most complex construct could be required (Gefen, Straub et al. 2000). Small sample size often renders the results of data analysis suspect. In terms of the statistical significance of the results, it is suggested that the practical significance of the results deserves more attentions (Lucas 1991). After data analysis, the researcher interprets the results, determines their implications, and writes up the results.

Survey research can be problematic, for example in the cases of:

- Trivial research questions (Lucas 1991) .
- Failure to use theory to guide the investigation (Lucas 1991).
- Poor models of inquiry (Lucas 1991; Lee, Barua et al. 1997).
- Weak hypotheses (Lucas 1991).
- Missing factors in analysis. ( Lee, Barua et al. 1997)
- Mistaken assumptions of the representative nature of samples, and
- The production of results that leave the researcher more confused than before (King 1991; Lucas 1991; Myers 1997; Lee 1999).

In addition, survey research may lack responsiveness to discoveries made during data collection. Accordingly, to improve the quality of survey research, Lucas (1991) suggested finding an interesting research problem, formulating a compelling model and hypotheses, and developing a rigorous research design to test the model. While large sample sizes enable the use of advanced statistics (Gefen, Straub et al. 2000), such as multivariate analysis procedures, which will lead to good statistical power (Baroudi and Orlikowski 1989), multiple sources of data increase the credibility of results.

For this research, the positivist perspective and survey research method is considered appropriate, for three reasons. These are

- Firstly, the positivist perspective and survey study is a dominant, well-established, and well-accepted research approach in information systems research (Kraemer 1991; Palvia, Mao et al. 2003; Westland 2004).
- Secondly, the nature of this study addresses hypothesis testing, and the results

are expected to be generalizable

- Thirdly, survey research has already been used in studies of adaptive structural theory and its applications (for example Gopal, Bostrom et al. 1993; DeSanctis and Poole 1994; Chin, Gopal et al. 1997; Salisbury, Chin et al. 2002).

Internet surveys are considered to be the most suitable means of data collection for this study. This is because

- Firstly, the study involves Internet and Internet-related technologies.
- Secondly, the target population for this study is internet-related technology and systems users who are assumed to be cross-sectional, scattered, and mobile.
- Thirdly, the research has a limited budget (refer to section 4.3.3 for details).

The following sections further discuss survey research-related methods and techniques, including the survey scale development process, Internet surveys, data analysis methods, the research design and the sample design.

### **4.3.2 The Survey Scale Development Process**

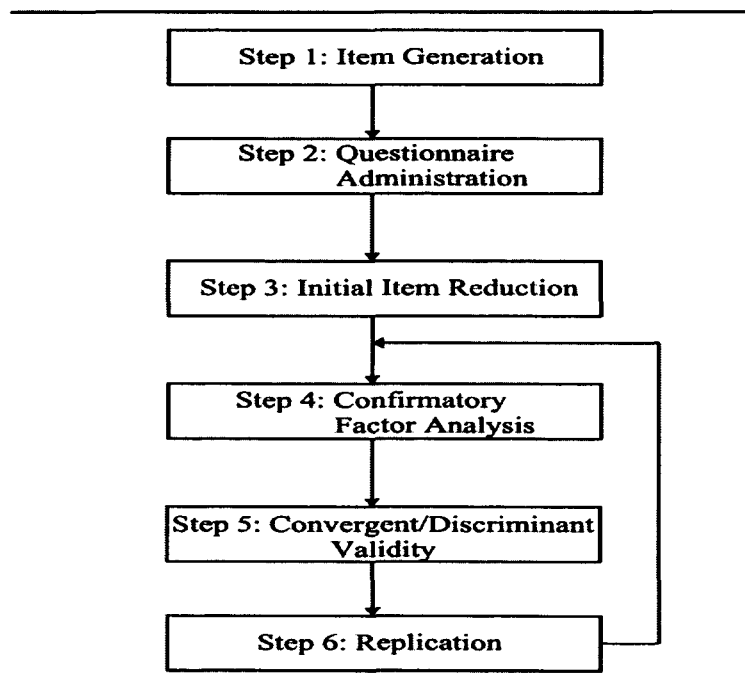
Ensuring that the survey scale actually measures what it is claimed to measure plays a crucial role in survey research. Flawed measures, with problems such as inappropriate domain sampling, poor factor structure, low internal consistency reliability, and lack of basic validity would result in the findings in this study being untrustworthy. The adequacy and accuracy of measurement of the constructs of interest, i.e., construct validity, still remains a great challenge for MIS study which uses survey research methodology (Straub 1989; Boudreau, Gefen et al. 2001; Straub, Boudreau et al. 2004).

Structuralized survey scale development processes have been proposed for guiding survey scale development so as to ensure that the survey scales adequately and accurately represent the research constructs (Churchill 1979; Straub 1989; Hinkin 1998). Following these well-established procedures, quality survey scales can be achieved.

Figure 4.1 exhibits a scale development process suggested by Hinkin (1998). This process consists of six major steps, in which each step contributes to increasing confidence in the quality of the new measure, a brief discussion on the development of a quality survey instrument, based on the process, is presented below.

The first step deals with generating items to represent (or measure) adequately the construct of interest. A well-articulated content domain of the construct under examination is the essential prerequisite to achieve this target. Two important techniques applicable of creating preliminary items are the deductive method and the inductive method (Hinkin 1998). The deductive method is theory-driven (Hox 1997), and considered suitable for this study. According to Hinkin (1998), the deductive approach, if properly conducted, will help to assure content validity in the final scales.





**Table 4.1 The Scale Development Process (Hinkin 1998)**

The deductive approach relies on the operational definition of the construct under examination, derived from the theoretical definition of this construct, to develop the initial set of items (Hinkin 1998). These items are the empirical indicators which should suitably represent the each of the dimensions of the construct (Hox 1997). Therefore, an understanding of the phenomenon to be investigated, and a well-articulated theoretical definition of the construct of interest are essential. The main methods for developing measurement items by the deductive approach include dimension/indicator analysis (Hox 1997), semantic analysis (Hox 1997), and facet design (Shye, Elizur et al. 1994). Moreover, a number of rules are available to guide the writing of items in order to assure scale quality (Churchill 1979; Sudman and Bradburn 1982; Hinkin 1998).

After an initial set of items has been produced, a content validation assessment may apply. The content validity refers to

*“The degree to which a measure’s items are a proper sample of the theoretical content domain of a construct” (Schriesheim, Powers et al. 1993, p.386).*

Several techniques can be used to assess the content validity, such as literature reviews (Straub, Boudreau et al. 2004) and the quantitative method suggested by Schriesheim,

Powers et al. (1993). However, as indicated by Straub, Boudreau et al. (2004), there lacks clear consensus on the methods and means of determining content validity. As a result, content validity assessment, although desirable, is still found to be infrequent in IS research (Straub, Boudreau et al. 2004).

The second step, questionnaire administration, conducts data collection from a sample representative of the actual population of interest using the new questionnaire. The sample will be used for a preliminary study of the new measurement scales.

At step three, the initial item reduction is achieved through corrected-item total correlation analysis (Doll and Torkzadeh 1998), followed by the internal consistency assessment with Cronbach's alpha (Churchill 1979; Hinkin 1998), and then exploratory factor analysis (EFA), which is used to identify items that are not factorially pure (Doll and Torkzadeh 1998).

The corrected-item total correlation analysis, internal consistency assessment, and EFA allow the reduction of a set of items (observed variables) to a smaller set of items. The corrected-item total correlation analysis enables cleaning of the scale items by removing items whose corrected-item total correlation is  $< 0.5$  (Doll and Torkzadeh 1998).

Reliability reflects the accuracy of the measuring instrument. The internal consistency assessment with Cronbach's alpha is recommended for assessing the reliability of the measurement scales (Churchill 1979; Hinkin 1998). As a rule of thumb, the acceptable Cronbach's alpha value could be as low as 0.5 or 0.60 for exploratory research, and 0.7 for confirmatory research (Churchill 1979; Straub, Boudreau et al. 2004).

The EFA is used for seeking a more parsimonious and interpretive factorial structure of the construct of interest, therefore, only those items that clearly load on a single appropriate factor should be retained (e.g., factor loadings of greater than 0.4) (Hair, Anderson et al. 1995). However, items that load on more than one factor, with cross loading of greater than 0.4 should be dropped.

The preliminary study results in a revision of the survey instrument, which is subjected to a further construct validation. The revised instrument is sent to a sample of the target

population to collect data. Construct validation is conducted with confirmatory analysis using the collected data. The aim of the confirmatory analysis is to completely validate those newly developed scales, or the scales adapted from previous research which are to be used for different purposes in different contexts. The primary activities during the confirmatory analysis involve assessment of convergent, discriminant, and nomological validity, with each successive validity measure providing a more rigorous test of the scales than the previous one (Straub 1989; Chin, Gopal et al. 1997). The confirmatory analysis is also referred to as measurement model testing when conducted using structural equation modeling techniques, such as covariance-based SEM (e.g., LISREL) and PLS (Joreskog and Sorbom 1996; Chin, Gopal et al. 1997; Straub, Boudreau et al. 2004).

As illustrated in (Chin, Gopal et al. 1997; Salisbury, Chin et al. 2002), practical assessment procedures of construct validity in IS studies usually use a systematic approach to establish each form of validity through a sequential step of tests, in the form of simple to advanced, i.e., the constructs are subjected to examine individually in isolation first, then in relation to other constructs, and finally in a nomological network (or research model). This is because the outcomes of measurements depend on not only the scales' individual characteristics, but their specific relationship with other constructs that also influence the outcomes. As indicated by Chin, Gopal et al. (1997), while the sequential tests are necessary to prove validity for a scale, only the final nomological test is sufficient to prove validity for a scale.

In organizational research, both Multitrait-Multimethod Matrix (MTMM) (Churchill 1979) and confirmatory factor analysis (CFA) (Bagozzi, Yi et al. 1991) are considered as major approaches for examining convergent and discriminant validity (Straub 1989; Hinkin 1998; Straub, Boudreau et al. 2004). As MTMM usually involves different traits and different measurement methods in the each of the traits, this may sometimes appear to be difficult in practice (Hinkin 1998). Thus the CFA technique is considered more applicable, and recommended for convergent and discriminant validity testing (Hinkin 1998).

Convergent validity refers to examining the extent to which a survey scale correlates with other measures designed to assess similar constructs (Straub 1989). If constructs

are valid in this sense, relatively high correlations between measures of same constructs using different methods can be expected. Ideally, maximally different methods are expected for examining the convergent validity (Goodhue 1998).

CFA for convergent validity works in the following way. The researcher specifies a factors model, which includes factors, measures (or items) to the factors, and the linkages among the factors based on previous studies or theory. CFA tests the fit of that model against the given data set, and determines how well the model explains the sample data (Chin, Gopal et al. 1997; Gefen, Straub et al. 2000). The convergent validity of a tested construct is established if all the item loadings are above 0.60, and the overall model goodness of fit indices are adequate (Chin, Gopal et al. 1997).

While the scales were tested in isolation for convergent validity, it is considered important to investigate construct validity by testing the scales in relation to other constructs of interest which are parts of the research model, to ensure that the items of the tested scales do not better relate to these other constructs (Chin, Gopal et al. 1997).

Discriminant validity is concerned with the extent to which the validated scales do not correlate with other dissimilar measures of interest (Churchill 1979; Goodhue 1998). This further testing on discriminant validity provides a more rigorous assessment of construct validity than the previous testing (e.g., convergent validity)(Chin, Gopal et al. 1997).

Discriminant validation can be performed by testing the correlation between constructs, or by comparing the Chi-Square ( $\chi^2$ ) value of two analyses (Chin, Gopal et al. 1997). If the correlation is not equal to 1.00 and the overall model fit indices are adequate, discriminant validity would be indicated. However, it can be more rigorously tested using  $\chi^2$  difference as the result of comparing the  $\chi^2$  measures for two analyses. In the first analysis, the constructs of interest are assumed to be identical, by fixing the correlation between the constructs at 1.00, and then calculating a  $\chi^2$  measure for the model. In the second analysis, the correlation between the constructs is released, i.e., allowed to be freely estimated. A new  $\chi^2$  can then be generated for the same model against the same dataset. In this test, there are no differences in the model specifications for the two analyses except for the correlation change. As a result, discriminant validity is suggested if the ratio of difference of the  $\chi^2$  values for two analyses to the difference

in the degrees of freedom between the two models is greater than 3.84 (1 d.f.,  $\alpha = 0.05$ ) (Chin, Gopal et al. 1997).

Nomological validity addresses the “behaviour” of the newly measured construct in an accepted network of theory (also called a nomological network or research model), which belongs to the network (Straub, Boudreau et al. 2004). If the newly measured construct fits the network, and acts in expected ways in relation to other well-established constructs, and the overall model goodness of fit indices are adequate (in a sense of SEM), the nomological validity of the scale instrument is suggested (Goodhue 1998; Straub, Boudreau et al. 2004). Therefore, the existence of an accepted nomological network is a prerequisite for the assessment of nomological validity. However, as recommended by Straub (1989) and Straub, Boudreau et al. (2004), previously validated instruments should be exploited and adapted wherever and whenever possible. For the adapted instruments, a further validation is necessary, even though the original instruments have previously been carefully validated (Straub 1989). In summary, a viable survey development process can be formulated for this study. As this research utilizes previous studies of knowledge management, structuration theory and adaptive structuration theory, and information systems success (see chapter II and III for more details), the operationalization of the research constructs are theory-driven. As a result, the measurement items for the research constructs will be generated in a deductive way. The content validity will be established by a suitable review of literature and constructs development.

A preliminary study, including data collection, internal consistency and reliability analysis, and exploratory factor analysis, will be conducted for pre-testing the psychometric properties of the newly developed scales, resulting in a revision of the initial survey scales. This will be followed by formal data collection using the revised survey scales.

With the data collected using the revised survey scales, the scales are subjected to a further constructs validation process. The convergent validity and discriminant validity are assessed through confirmatory factor analysis (CFA), and nomological validity is achieved through the model testing processes, both of CFA and model testing are conducted using structural equation modeling techniques, so as to assure the good

psychometric properties of the measurement scales before testing the hypotheses.

The next section will focus on the Internet surveys, which are the main data collection method for this research.

### **4.3.3 Internet Surveys**

Internet surveys are rapidly becoming an important means of data collection in organizational studies due to the universal use of the Internet (Zikmund 2000; Simsek and Veiga 2001). According to Simsek and Veiga (2001), there exist two modalities of Internet surveys, namely the email-based survey and the Web-based survey. The email survey is a computerized, self-administered questionnaire, which is sent, received, completed, and returned through email systems. Email surveys can be classified into three types. The first type includes the questionnaire as part of the message text, the second involves the questionnaire as an attachment to an email, and the third emails the recipient an email with a URL embedded in the email text, which directs the recipient to a web site to download the survey questionnaire, complete, and then return it via email. The web-based survey involves a computerized, self-administered survey questionnaire hosted by a web site, which is accessed and completed by respondents via compatible web browsers. Respondents can be diverted to the web-based survey questionnaire by various means including invitation emails, information pages on web sites, or postal notifications. Web-based surveys are becoming the dominant forms of research data collection due to their benefits in saving respondents' time, saving costs, and speeding up the data collection process (Zikmund 2000).

Internet self-administered surveys have strengths compared with conventional survey methods. They have the potential of radically reducing the cost of conducting surveys, for example,

- No paper is required for the survey.
- Data can be transferred directly from the form into the statistic analysis software.
- No additional costs as the size of sample increased.
- Fast data collection ( Simsek and Veiga 2001).
- They can reach a large group of people rapidly and immediately regardless of

their location which results in a shorter data collection process and substantial time saving (Zikmund 2000).

The Internet survey process is easy to manage. For example, a follow-up questionnaire or reminder can be sent to potential respondents easily and efficiently, which may improve the response rate (Simsek and Veiga 2001). In addition, web technologies can help to improve the response quality and quantity of surveys, for instance, the questionnaire can be designed so that questions can be adapted according to the respondent's answer to the previous questions. Text, sound, graphics, and live interaction can provide a richer medium, to stimulate potential respondents' willingness to participate in the survey.

The major concerns raised in Internet self-administered surveys involve:

- Representativeness of the sample.
- Sampling frame.
- Sampling control.
- Non-response error.
- Measurement error, and
- Anonymity and confidentiality (Simsek and Veiga 2001).

Although these also exist in traditional pencil-paper self-administered surveys, they are emphasized by electronic communications and web technologies. While the Internet survey can only be conducted with those can and do use the Internet, concerns over Internet self-administered surveys will remain.

The sampling representativeness may not be a problem when the target population is Internet users, or target organizations who have widespread access and use of Internet among their staff. One way to improve sampling representativeness is the simultaneous use of multiple survey methods, that is, Internet surveys used in combination with other data collection techniques, such as postal surveys and telephone interviews, so as to reduce coverage error (Zikmund 2000; Simsek and Veiga 2001).

The concerns over the sampling frame for Internet surveys arise when the survey

involves cross-sectional sampling. This is because either the Internet lacks universal coverage in many populations or because of a lack of available lists of potential respondents. While it is critical to construct reasonable and usable sampling frames for Internet self-administered surveys, several methods can be used to achieve the target. For example, listservs, discussion forums, and online public directories could be useful resources for culling sampling frames ( Simsek and Veiga 2001).

Sampling control could be achieved to a certain extent by sending the survey or an invitation to previously identified individuals. Access control can be used to reduce the possibility of multiple responses from a single individual.

For any survey, the response rate is a concern. Non-response error occurs when some potential respondents of the targeted sample do not respond to the survey. Empirical evidence shows that the response rates for Internet survey range from 7% to 76% (Simsek and Veiga 2001). Although there are few theoretical examinations of non-responses in Internet surveys, several ways of improving response rate in Internet surveys have been suggested (Stanton 1998; Simsek and Veiga 2001). For example, adding an introduction to the Internet survey, which includes:

- A description of the reasons for the survey.
- An explanation of the worthiness of the survey.
- An appeal that stresses the potential respondent's importance in participating in the survey, and
- An appeal to help the researcher in completing an important project.is believed to be helpful in increasing response rate (Simsek and Veiga 2001).

In addition, the articulation of sponsorship and follow-up emails are considered as effective strategies of increasing response. As noted by Simsek and Veiga (2001), the Internet survey questionnaire should be kept as simple and short as possible so as to save potential respondents' time and online access costs.

In terms of measurement error, especially systematic measurement error, an empirical investigation shows little difference between the Internet survey and traditional self-administered survey methods (Stanton 1998). However, the degree of anonymity in



Internet survey may affect the way of respondents answering survey questions, as the responses may suspect that they can be tracked by means of email or web browser monitoring systems or by “cookies”. This will cause systematic measurement errors in the data collected (Simsek and Veiga 2001).

While complete anonymity and confidentiality of responses are not possible with Internet survey, managing anonymity and confidentiality effectively will not only increase response rates, but also improve the quality of data collected by reducing systematic measurement errors. Greater anonymity will result in greater confidentiality in respondents. Effective ways of managing anonymity and confidentiality will render the respondents completely untraceable and unidentifiable. A set of means of support of anonymity and confidentiality has been suggested (Simsek and Veiga 2001), which includes submitting the completed surveys via anonymous re-mailers, stopping the use of cookies in web surveys, or setting up the survey questionnaire on an alternative outside web site. However, potential respondents should be informed that the nature and degree of anonymity and confidentiality afforded, the voluntariness of participation in the research, the research specification and purpose, the security of documents or records of data obtained in the research, how the data will be used and who will have access to it. While potential respondents are likely to believe that Internet surveys lack confidentiality and anonymity, disclosure of this information may help the researcher to gain the trust of respondents.

In summary, Internet self-administered surveys offer researchers exciting new possibilities for data collection. Internet surveys provide a good opportunity for researchers who have a limited research budget, who are interested in fast data collection, or who are doing research on internet-related technologies and systems, in which the target populations are Internet users who may be scattered or mobile. However, the full potential of Internet surveys is at present impeded by the Internet’s lack of universal coverage, difficulties in constructing reasonable and usable sampling frames and sampling control, and lack of respondent anonymity and confidentiality. Hence, it is crucial to utilize Internet surveys only for appropriate research projects.

In this section, a great deal of attention has been given to the Internet survey method and process, which serves as the main data-collection tool for this study.

As recommended by Gefen, Straub et al. (2000), second-generation data analysis techniques, i.e., Structural Equation Modeling (SEM), should be used extensively in IS research in order to achieve high quality statistical analysis. SEM has been chosen as the statistical analysis technique for this study. The next section reviews SEM techniques, which include covariance-based SEM (e.g., LISREL) and partial-least-squares-based (PLS-based) SEM.

#### **4.3.4 Structural Equation Modelling (SEM)**

Structural equation modelling (SEM) is recommended as an important second generation statistical technique for high quality information systems research (Gefen, Straub et al. 2000; Boudreau, Gefen et al. 2001; Chin, Marcolin et al. 2003). SEM differs greatly from most of the first generation regression models, such as linear regression, ANOVA and MANOVA. SEM allows multiple layers of linkages between independent, mediating, and dependent variables to be expressed and analyzed through hierarchical or non-hierarchical, recursive or non-recursive structural equations to present a more complete picture of the entire model (Gefen, Straub et al. 2000). When the phenomena of interest are complex and multidimensional, SEM is the only analysis that allows complete and simultaneous tests of all hypothesized relationships. It can, thus, provide complete set of statistic analysis results simultaneously in a single, unified process. This is unlike first generation statistic techniques which can only support single and isolated statistic analyses, and thus show an incomplete picture of the entire model (Tabachnick and Fidell 1996; Chin and Newsted 1999; Gefen, Straub et al. 2000).

SEM comprises two interrelated models,

- The structural mode, a set of hypothesized dependence relationships among the model constructs, and
- The measurement model, which specifies indicators (observed variables) for each of the model constructs.

SEM enables researchers to assess the structural model and measurement model simultaneously (i.e., a holistic analysis). The result is a more rigorous analysis of the proposed research model because measurement errors of the observed variables are analyzed as an integral part of the model, and factor analysis is combined in one operation with the structural model testing /hypotheses testing. Thus, SEM provides the capabilities of testing both the measurement properties of instruments and the hypotheses in the same analysis (Kelloway 1998; Gefen, Straub et al. 2000). In IS studies (e.g., Fichman and Kemerer 1997, Teo, Wei et al. 2003), data analysis using SEM usually exploits a two-phase strategy. The first phase focuses on the assessment of measurement model so as to assure the psychometric properties of the research constructs, and, in the second phase, the structural model is evaluated to test the

hypothesized relationships.

SEM involves a set of specific statistical terms, including:

Latent variable/construct (LV). This refers to a research construct that cannot be measured directly but rather assessed indirectly using a number of observable variables (indicators). There exist two distinct measurement modes, i.e., reflective mode and formative mode, representing different relationships between indicators and their associated latent constructs in SEM (Gefen, Straub et al. 2000).

Reflective variable. Under the reflective mode, indicators represent and reflect the latent variable, and the indicators are characterized as unidimensional and correlated.

Formative variable. Under the formative mode, indicators cause and form the latent variable, that is represent different dimensions of the latent variables, and these indicators are not assumed to be necessarily correlated with each other.

Exogenous and endogenous latent constructs. An exogenous latent construct refers to a construct that acts only as an independent variable, i.e., the predictor or “cause” for other variables in the model, while the endogenous latent construct is the dependent variable. In SEM path diagrams, there are always one or more arrows leading out of exogenous latent constructs and into endogenous latent constructs.

Item loadings. Item loadings provided by SEM represent the degree of correlation between the observable indicators (items) and derived factors.

Structural relationships and Path coefficients. Structural relationships and standardized path coefficients refer to the hypothesized relationships between research constructs under investigation, and the standardized path coefficients indicate the relative strength of the statistical relationships which can be interpreted in a very similar manner to regression coefficients (Gefen, Straub et al. 2000).

There are two distinct classes of SEM techniques, resulting from the use of different algorithms for analysis, namely covariance-based SEM and partial-least-squares-based

SEM (PLS-based SEM) (Kelloway 1998; Chin and Newsted 1999; Gefen, Straub et al. 2000; Chin, Marcolin et al. 2003). While covariance-based SEM is designed to minimize the difference between the sample covariances and those predicted by the proposed model in terms of model fit indices and residuals, PLS-based SEM, as a variance-based approach, is designed to maximize the explained variance in all endogenous (dependent) constructs, similar to examining the corresponding  $R^2$  (R-square) values in linear regression. The difference between these two SEM techniques lies in the objective of their analysis, their usages, the statistical assumptions on which they are based and the nature of the model fit statistics they produce (Gefen, Straub et al. 2000; Chin, Marcolin et al. 2003).

While the objective of covariance-based SEM is to show that the operationalization of the theory being examined is corroborated and not disconfirmed by the sample data, the objective of the PLS-based SEM is to reject a set of path-specific null hypothesis of no-effect by showing high  $R^2$  (R-square) and significant t-values (Gefen, Straub et al. 2000).

In terms of respective usage, the covariance-based SEM usually requires a sound theory base and is best suited for confirmatory research, whereas PLS-based SEM does not necessary require a theory base and supports both exploratory and confirmatory research (Chin and Newsted 1999; Gefen, Straub et al. 2000).

Covariance-based SEM usually requires multivariate normal distribution of measured variables when model estimation is through maximum likelihood (ML) and generalized least squares (GLS) (Gefen, Straub et al. 2000). The multivariate normality assumes that each variable and all linear combinations of the variables are normally distributed, and can be examined partially through normality, linearity, and homoscedasticity of variables (Tabachnick and Fidell 1996). However, the check is considered as a necessary, but not a sufficient condition for ensuring multivariate normality (Tabachnick and Fidell 1996).

As the multivariate normality of measured variables is so difficult to ensure in practice, skewness and kurtosis have been suggested to justify the applicable of covariance-based SEM techniques, such as LISREL (West, Finch et al. 1995; Tabachnick and Fidell 1996). As a rule of thumb, West, Finch et al. (1995) suggest if skewness  $<2$  and kurtosis

<7, the underlying distributions of measured variables can be considered not to be substantially non-normal, therefore LISREL is applicable in this case.

While the assumption of multivariate normality could ensure valid statistical analysis, multivariate statistics may sometimes be robust to violation of their assumptions for covariance-based SEM (Tabachnick and Fidell 1996). Similarly, Gefen, Straub et al. (2000) suggests that ML and Generalized Least Squares (GLS) are applicable even when the measured variables deviate from the assumption of multivariate normality. Moreover, if there is a large sample (e.g., 200 and more), but the researcher is not able to assume multivariate normality, GLS estimation is suggested to be used (Kelloway 1998). In addition, when there are substantial deviations from multivariate normality, Weighted Least Squares (WLS) is the estimation method of choice (Joreskog and Sorbom 1996; Gefen, Straub et al. 2000).

PLS-based SEM, however, is relatively robust to deviations from a multivariate distribution (Chin, Marcolin et al. 2003). Consequently, the examination of statistical assumptions, such as multivariate normality, may not be necessary in applying PLS. Furthermore, both the covariance-based and PLS SEM assume linear relationships between the observed variables and their constructs, and between one construct and another (Gefen, Straub et al. 2000).

For the model fit statistics in the covariance-based SEM, the most important and widely used entire model fit indices include the likelihood-ratio chi-square ( $\chi^2$ ), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), and Root Mean Residual (RMR) (Tabachnick and Fidell 1996; Gefen, Straub et al. 2000). In the case of PLS-based SEM, however, good model fit is confirmed with significant path coefficients, i.e., acceptably high  $R^2$  and construct reliability and validity for each construct (Hulland 1999; Gefen, Straub et al. 2000; Chin, Marcolin et al. 2003; Gefen and Straub 2005).

While PLS-based SEM is applicable in the case of a small sample (Chin and Newsted 1999), covariance-based SEM is considered as a large sample technique for two reasons (Tabachnick and Fidell 1996; Kelloway 1998). Firstly, the Covariance-based SEM is based on covariances, which are less stable when estimated from small samples; secondly, both the estimation methods (e.g., maximum likelihood, ML) and tests of

model fit (e.g.,  $\chi^2$  test) are very sensitive to sample size, and are based on the assumption of large samples (Gefen, Straub et al. 2000).

There is a variety of recommendations on the minimum sample size required for statistical analysis in both Covariance-based SEM and PLS. For example, Gefen, Straub et al. (2000) suggested that a sample of 100-150 is essential for a plausible analysis using covariance-based SEM, depending on the degree of model complexity. Although a minimum sample size of 100 still enables a statistical analysis by Covariance-based SEM, it has been recommended that a sample of 200 or more results in more accurate analysis and reliable conclusions (Kelloway 1998). Alternatively, the required sample size can also be worked out according to a ratio suggested by Tabachnick and Fidell (1996), i.e., the ratio of sample size to estimated parameters could be between 5:1 and 10:1. The minimal sample size required for PLS-based SEM depends on the most complex construct in the model, and at least 10 times the number of items in the construct is necessary (Chin and Newsted 1999). Therefore, for PLS-based SEM, the minimal recommendations could range from 30 – 100 cases (Chin and Newsted 1999; Gefen, Straub et al. 2000). As a result, PLS-based SEM is more suitable for the analysis of small data samples, which may not exhibit multivariate normal distribution (Chin and Newsted 1999; Gefen, Straub et al. 2000; Chin, Marcolin et al. 2003).

Another important difference between the two SEM techniques is that only covariance-based SEM can provide the capability of assessing uni-dimensionality, i.e., the extent to which items load only on their respective constructs. Thus, covariance-based SEM is considered to provide better coefficient estimates and more accurate model analyses when the research model has a sound theoretical base and the model variables have measurement errors.

In addition, PLS-based SEM is recognized as the only available second-generation statistic analysis technique applicable to dealing with measurement models with both formative and reflective indicators (Chin and Newsted 1999; Chin, Marcolin et al. 2003), whereas the measurement models for co-variance-based SEM can only comprise reflective indicators (Gefen, Straub et al. 2000).

Strategically, PLS-based SEM, more suited for predictive applications and theory

building, is suggested to be a complementary technique and a forerunner to the more rigorous covariance-based SEM, which is best suited for confirmatory research (Gefen, Straub et al. 2000; Chin, Marcolin et al. 2003).

The application of PLS in a management research context usually involves three sets of analyses (Hulland 1999):

- Assessment of the adequacy of the measurement model by examining measurement reliability and validities.
- Assessment of structural relationships by checking path coefficients and respective t-values, and corresponding R-square values for the endogenous constructs.
- Interpretation of path coefficients and determination of model adequacy.

For practical purposes, some rules of thumb thresholds have been established for justifying PLS statistics in management research. For instance, the adequacy of a measurement model can be established if measurement item loadings are above 0.7 (Hulland 1999), composite reliabilities are above 0.70 (Nunnally 1978), the AVE (Average Variance Extracted) of each construct is larger than its correlation with any other constructs in the model, and each item has a higher loading on its assigned construct than on any other constructs (Gefen, Straub et al. 2000). Furthermore, good model fit also requires significant path coefficients, and acceptably high  $R^2$  (Gefen, Straub et al. 2000; Gefen and Straub 2005).

SmartPLS (currently in its 1.01 version) is a software package for applying PLS-based SEM on a graphical user interface, developed by the Institute for Operations Management and Organization at the University of Hamburg, Germany in 2004 (Hansmann and Ringle 2004).

SmartPLS generates a set of statistics. For assessing the adequacy of the measurement model, SmartPLS provides item loadings and significance level (t-values), composite reliabilities and AVE (Average Variance Extracted) for each of latent variables, which can be used for examining the reliability and validity of measures. The reliability and validity of measures include individual item reliabilities (i.e., using item loadings),



convergent validity (i.e., using composite reliability), and discriminant validity (i.e., using AVE and the correlation coefficients between constructs) (Hulland 1999; Gefen, Straub et al. 2000; Straub, Boudreau et al. 2004).

At the structural model level, SmartPLS produces path coefficients with t-values, i.e., the strength and significance of hypothesized relationships, and correlations among the latent variables, together with the individual  $R^2$  for endogenous constructs. The T-values of both path coefficients and item loadings are calculated using the bootstrap method in SmartPLS (Yung and Chan 1999).

As mentioned above, covariance-based SEM should be mainly used as a confirmatory analysis method for the validation of instruments and for theory testing. LISREL is a computer program (currently in its eighth version) that performs covariance-based structural equation modelling (Joreskog and Sorbom 1996; Kelloway 1998). In terms of instrument validation, LISREL enables the assessment of construct validity, consisting of convergent validity and discriminant validity, and analyzing constructs reliability. LISREL provides the facility for testing hypotheses through evaluating the overall fit of the observed matrix with the hypothesized covariance model. LISREL generates three sets of statistics at three levels (Gefen, Straub et al. 2000).

At the individual path and construct level, the most important statistics include item loadings, measurement error along with their respective t-values, the derived construct reliability, and the Squared Multiple Correlation (SMC) of each of the exogenous latent constructs.

At the overall model fit level, the most important statistics comprise the likelihood-ratio chi-square ( $\chi^2$ ), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Root Mean Residual (RMR), and Normed Fix Index (NFI).

The third set of statistics addresses the individual path modification indexes, which estimate the difference in model fit  $\chi^2$  for each possible individual additional path.

For practical purposes, some rules of thumb thresholds have been established for justifying the above statistics in information systems research (Gefen, Straub et al.

2000). For example, the minimum level of loading is 0.60 for convergent validity (Chin, Gopal et al. 1997); the construct reliability should be above 0.70 (Gefen, Straub et al. 2000); the  $\chi^2$  statistic should be insignificant with a p-value above 0.05 or a  $\chi^2$  as small as possible or the ratio of the  $\chi^2$  to degree of freedom should be smaller than 3:1 (Gefen, Straub et al. 2000), or 5:1 (Salisbury, Chin et al. 2002); the GFI should be above 0.90, the AGFI should be above 0.80, and RMR should be below 0.05 (Gefen, Straub et al. 2000), or 0.10 (Salisbury, Chin et al. 2002); NFI should be above 0.90 (Chin and Todd 1995).

While covariance-based SEM techniques such as LISREL8 are considered as a state-of-the-art, powerful second-generation statistics technique for confirmatory research (Gefen, Straub et al. 2000), they should, however, be utilized with caution. Firstly, covariance-based SEM technique should only be used as a confirmatory method, which means that the researcher's model must be based on a solid theory base. Secondly, the researcher's model needs to be linear, i.e., the relationship between observed variables and their constructs and between one construct and another is linear. Thirdly, attention should be paid to carefully checking sample size and data distribution before conducting a formal statistic analysis using covariance-based SEM technique. Fourthly, it is crucial to ensure that the values of appropriate statistics reported are within the established thresholds, so as to make the research significant in the sense of good positivist science.

#### **4.3.5 Sampling Design and the Unit of Analysis**

The purposes of this research are to empirically investigate why users accept and use KMS, and the impacts of KMS on individual's social capital within knowledge organizations as well (organizational social capital can be formed partially by aggregating individual's). The target population for this study is knowledge workers, who are the users of KMS and working for a variety of different organizations (Sveiby and Lioryd 1987; Drucker 1999). In survey research, there are two typical sampling designs, i.e., cross-sectional design and longitudinal design (Vitalari 1991). According to Vitalari (1991), the longitudinal design usually refers to collect research data from the same organization several times in a given period of time (e.g., two times in a year) aiming to capture the change of the studied phenomenon over time, while the cross-sectional sampling design refers to collect data from a variety of organizations at the same time

providing a snapshot of the studied phenomenon across a variety of organizations at a certain time. The cross-sectional sampling design is recognized as primary method of research data collection (Lucas 1991). As the most popular data collection method in quantitative IS research, cross-sectional sampling design is considered suitable for this study because generalizability is important for this research, and a range of KMS and organizations are expected to be involved in the study.

The unit of analysis should fit the theoretical questions under investigation and the theoretical frame used to interpret data (Gutek 1991; Kling 1991). The individual has thus been chosen as the unit of analysis because the research interest is in both the acceptance and use of KMS, and the impacts of KMS use on user's social capital suggest conducting data collection and analysis on the individual level.

#### **4.3.6 Summary**

Following the theoretical discussion about three popular research approaches in management study in the previous section, this section focuses on the positivist methodology and the methods applicable to this research study.

A viable survey scales development process for this research has been formulated in subsection 4.3.2, based on a combination of the survey scales development process developed by Hinkin (1998), the scales development processes designed by Churchill (1979), and the scale development process proposed by Straub (1989).

An Internet survey has been chosen as the primary data collection method for this research because the target population of this study is the users of Internet-related technology systems (i.e., KMS). Furthermore, generalizability is important for this research, and a range of KMS and organizations are expected to be involved in this study. Subsection 4.3.3 presents a detailed discussion of Internet surveys. Subsection 4.3.4 focuses on a comprehensive introduction about the structural equation modeling techniques (PLS and LISREL), which will be exploited to conduct the survey scales development and hypothesis testing in this specific study. The sampling design and unit of analysis are mentioned in Subsection 4.3.5.

In sum, this section focused on the methods issues related to this research, and has

articulated methodological principles and processes for conducting the research.

The remaining part of this section will pursue the development by presenting processes of scales development, a preliminary study of the scales, revision of the scales, and a formal survey using the revised instrument.

## 4.4 The development of the Survey Questionnaires

The AST-based KMS success model (Figure 3.10), presented in Section 3.8 of Chapter 3, involves two major sets of variable. The independent variables, also referred as appropriation of KMS, include:

- Performance-related (or function-related) use of KMS,
- Information Quality/Task-Technology Fit,
- Social Norms,
- Perceived Usefulness, and
- Ease of Use
- The dependent variable is social capital development.

In addition, the research model also addresses a set of moderating variables, such as gender, age, and organization size. According to the literature review (refer to Chapter III), the dimensions and working definitions of these constructs are summarized in Table 4.2.

Dimensions	Working definitions
Perceived Usefulness (PU)	"The degree to which a person believes that using a particular system would enhance his or her job performance " (Davis 1989)
Perceived Ease of Use (EOU)	"The degree to which a person believes that using a particular system would be free of effort " (Davis 1989)
Performance-related Use of KMS (KMS-Use)	The extent to that a KMS is used to carry out particular tasks (e.g., information, communication, and collaboration).
Information Quality /TTF (INFOQ)	The degree to which a KMS meets the information needs of its end-users in terms of information content quality and information services quality (Goodhue 1998).
Social Norms (NORMS)	The extent to which individuals perceive support from management and peers in using a KMS, and desire to please management and peers by using the KMS (Lucas and Spitler 1999).
Social Capital Development	The degree to which a person believes that using a KMS

(SC)	would affect his/her social capital.
Function-related Use of KMS (Use-KMS)	Self-reported of the extent to that a selected set of KMS functions is used in people's daily work.

**Table 4.2 The Dimensions and Working Definitions of Appropriation of KMS**

Among the research constructs presented in Table 4.2, there exist well-accepted scales for “perceived usefulness (PU)”, “perceived ease of use (EOU)” (Davis 1989; Straub, Limayem et al. 1995), and “social norms (NORMS)” (Lucas and Spitler 1999). As these were not developed for KMS, they are subject to adaptation in wording and validation to the context of KMS. Accordingly, survey instrument development needs to focus on the scales for the performance-related use of KMS (KMS-Use), function-related use of KMS, TTF/Information Quality (InfoQ), and social capital development (SC).

As a key step before working out the measurement items, the operationalization, i.e., translating the theoretical construct into observable variables, of these research constructs need be carried out, followed by the development of initial measurement scales for these constructs.

#### **4.4.1 Operationalization of the Performance-related Use of KMS**

Operationalization of the performance-related use of KMS starts by identifying typical activities that are relevant to the two dimensions of the performance-related use of KMS: information-related use of KMS and interaction-related use of KMS. Typical activities relevant to the information-related use of KMS can be approached from three different perspectives. From the perspective of organizational knowledge management, information-related use of KMS is mainly relevant to the codification knowledge management, which emphasizes the production, storage, and utilization of explicit knowledge (Hansen, Nohria et al. 1999; Zack 1999). For instance, according to Borghoff and Pareschi (1998), in the Xerox KM architecture, the knowledge repositories allow users to search, publish, and access stored information, the technologies for the flow of knowledge enable users to efficiently distribute information/explicit knowledge to people, and the knowledge cartography provide users with the maps of expertise, experts, working experience, and interests.

From the perspective of supporting community of practice, according to Wenger (2001), typical activities relevant to the information-related use of KMS involve information distribution (e.g., news, reminders, questions, and so forth), messaging, accessing repositories for artefacts (e.g., best practices, war stories, discussion records, documents, tools), links to experts, and use of member directories with personal profiles.

From the knowledge work perspective, the information-related use of KMS enables knowledge workers to access and exchange information and explicit knowledge efficiently (Schultze 2003).

As this study is focused on the flow of knowledge within organizations, selected typical activities related to the information-related use of KMS include access to the information and knowledge repository, distribution of information and explicit knowledge and the location of people for information and knowledge.

Similarly, typical activities relevant to the interaction-related use of KMS can also be aggregated from the three perspectives. From the organizational knowledge management perspective, the role of KMS in personalization strategy is mainly concerned with enabling people to 'talk' with each other and supporting people-to-people knowledge transfer (Hansen, Nohria et al. 1999). For instance, the technical architecture for Xerox KM provides users with shared explicit knowledge and experience, a shared workspace and collaborative opportunities (Borghoff and Pareschi 1998).

From the perspective of community of practice, Wenger (2001) summarized the role of KMS in facilitating and supporting people interaction, such as engaging in communities, participating in online discussion group, discussing ideas, exchanging views, sharing experience, and collaborative and cooperative problem-solving (Wenger 2001).

From the knowledge work perspective, the interaction-related use of KMS mainly involves activities relating to problem solving within non-routine tasks (Elkjaer 2000). For instance, in Seven-Eleven Japan, information technologies were used to (Nonaka and Reinmoeller 2000) :

- Support dialogue and improvisation.
- Share problems and solutions.
- Assist the cross-boundary teams working in strategic alliances, developing new products with manufacturers, and
- Improve and speed up the process of collaborative product development and delivery.
- 

In summary, typical activities related to the interaction-related use of KMS include:

- Maintaining communication.
- Participating in online discussion groups.
- Working in virtual communities.
- Discussing/exchanging ideas/views/experiences, and
- Collaborating and cooperating with peers.

To sum up the above discussion, a summary of the operationalization of the Information-related use of KMS and interaction-related use of KMS can be found in Table 4.3.

Dimension	Working Definition
Information-related	The extent to which KMS is used to access information/knowledge repositories, distribute information, and identify and locate people for information and expertise.
Interaction (Communication & Collaboration)-related	The extent to which KMS is used to maintain communications, participate online discussion groups, engage in virtual communities, discuss/exchange ideas/views/experiences, and collaborate/cooperate with peers.

**Table 4.3 The Components and Working Definitions of KMS-use Construct**

It is worth noting that the information-related use of KMS represents one facet or dimension of the actual use of KMS in organizations, as does the interaction-related use of KMS. They can be, but are not necessarily correlated with each other. For instance, some KMS are designed mainly for sharing information and explicit knowledge through knowledge bases, such as a repository of best practice (O'Dell and Grayson 1998), and other KMS are mainly for supporting interaction activities (Wenger 2001). As a result, the performance-related use of KMS construct can be viewed as formed by the two sub-constructs. Furthermore, the performance-related use of KMS construct can be

operationalized as a formative, emergent construct formed from two sub-constructs: the information-related use of KMS and the interaction-related use of KMS.

**4.4.2 The Operationalization of Information Quality**

In the studies of information systems success, information quality is usually measured from a perspective of the user of the information by a set of information content attributes, such as accuracy, precision, currency, timeliness, reliability, completeness, conciseness, relevance, sufficiency, understandability and comparability (Bailey and Pearson 1983; King and Epstein 1983; Doll and Torkzadeh 1988; DeLone and McLean 2003). However, information services quality factors, such as locatability, accessibility, and assistance in using information (Goodhue 1998), are also crucial for information seekers. Accordingly, the information quality construct is conceptualized and operationalized as presented in Table 4.4.

Dimension	Working definition
Information Content Quality (labelled as Info_C)	The degree to which a KMS meets the information needs of its end-users in terms of information content quality, such as its accuracy, currency, and the right level of detail (Goodhue 1998).
Information Services Quality (Labeled as Info_S)	The degree to which a KMS meets the information needs of its end-users in terms of information services quality, such as locatability, accessibility, and assistance (Goodhue 1998).

**Table 4.4 The Dimensions and Working Definitions of Information Quality**

Conceptually, the two sub-constructs could be distinct, implying that information content quality is not necessarily correlated with information services quality. For instance, the expert locator in KMS helps a user find sources of information, but cannot guarantee the information content quality. Consequently, the information quality construct can be viewed as being formed by the two sub-constructs. In other words, the information quality construct can be operationalized as a formative, emergent construct formed from two sub-constructs: information content quality, and information services quality.

**4.4.3 The Operationalization of Social Capital Development**

Nahapiet and Ghoshal (1998) formulate the social capital construct with three major facets, namely a structural dimension, a relational dimension and a cognitive dimension.



The structural dimension is used to describe the overall pattern of connections between actors, i.e., the social networks (connections) around actors. The relational dimension reflects those assets created and leveraged through social networks. The cognitive dimension refers to a set of network-based resources providing shared interpretations, representations, and systems of meaning among actors. This definition of social capital focuses on those aspects of social capital that could facilitate the creation of intellectual capital (Nahapiet and Ghoshal 1998).

In an empirical study exploring the impact of social capital on value creation, Tsai and Ghoshal (1998) operationally defined the social capital construct at the level of business unit. In their study,

- The structural dimension of social capital was measured by social interaction ties between different units.
- The relational dimension was measured by perceived level of trust and trustworthiness between different units.
- The cognitive dimension was represented by perceived level of shared vision in the different units.

Social capital can be understood as an organizational phenomenon rooted in organizational fertility. The empirical study conducted by Tsai and Ghoshal (1998) can be considered as downstream research, which mainly focused on examining the impact of social capital on organizational performance, such as innovation and value. However, as the result of organizational life and individual interactions, the formation and development of social capital deserves further investigation. According to Cohen and Prusak (2001), social capital is crucial to organizational knowledge management, innovation, and productivity. In this study, therefore, the impact of KMS use on individuals' social capital development is examined, i.e., the degree to which using KMS would affect each component of individual's social capital.

Bearing the above social capital definition and measures in mind, working definitions for the impact of KMS use on each component of social capital are operationally defined as in Table 4.5

<b>Social Networks</b>	The degree to which a person believes that using KMS would enhance his or her social connections (e.g., know more people, making more acquaintances and friends, enhancing close relationship with colleagues).
<b>Trust building</b>	The degree to which a person believes that using KMS would enhance the level of trust between colleagues.
<b>Shared Vision</b>	The degree to which a person believes that using KMS would enhance the level of shared vision (e.g., organizational ambitions, vision, collective goals, and missions) with colleagues.

**Table 4.5 Components and Working Definitions of Social Capital Development**

The above three sub-constructs represent the three dimensions of social capital, and reflect the different aspects of social capital. Conceptually, the three sub-constructs form the social capital construct, and are not necessarily correlated with each other. Therefore the social capital construct can be operationalized as a formative, emergent construct formed by three sub-constructs: social networks, trust building, and shared vision.

#### **4.4.4 Positioning the KMS-related Instruments**

As a KMS is comprised of a set of loosely bundled capabilities and a mixture of multiple technical tools, KMS applications vary across organizations (Davenport and Prusak 1998). The questionnaire designed for investigating KMS applications in organizations should, therefore, evaluate the overall KMS applications and relevant services provided in organizations, not individual technology or application systems. In addition, the use of KMS is considered to be voluntary, even though employees in many organizations may be encouraged to use the technology as much as possible.

The end-users of the KMS applications are assumed mainly to be knowledge workers. A major purpose of the deployment of KMS is to facilitate and support organizational members in knowledge creation and sharing (Nielsen and Ciabuschi 2003), which mainly benefits knowledge workers. The questionnaire should thus be designed to target the individual user of KMS applications.

As the aim of the questionnaire is cross-sectional data collection involving a variety of organizations and KMS applications, it felt necessary from a practical point of view to clearly define KMS by identifying the major technical components; this was to help respondents answer the questions in the data collection processes. The major components of KMS are considered to be email, video conferencing, document/knowledge repositories, information/knowledge distribution, location of experts, online discussion forums, virtual communities, and virtual collaborative workplaces/teams (see Section 2.5.3), and all of these are included in the survey questionnaire.

All the variables with descriptions addressed in this study are summarized in Table 4.6.

Variable	Scale/Variable	Description
Performance-related Use of KMS Information-related usage Interaction-related usage	KMS-use Info_use CC_use	Use of KMS in obtaining information, sharing information, communicating with colleagues, and cooperating and collaborating with colleagues. KMS-use consists of two dimensions: information-related use (Info_use) and communication & collaboration-related use (CC_use).
Information Quality Information Content Information Services	INFOQ Info_c Info_s	Perceived information quality in terms of its accuracy, currency, the right level of detail, locatability, accessibility, and assistance. Information quality consists of two dimensions: information content-related, such as its accuracy, currency, the right level of detail, and information services-related, i.e., locatability, accessibility, and assistance.
Perceived Usefulness	PU	KMS improves effectiveness, and is very important in performing job.
Perceived Ease of Use	EOU	Ease to get KMS to do what is wanted, feeling very comfortable in using KMS.
Social Norms	Norms	Perceived support by management and peers in use KMS, and desire to please them by using KMS
Social Capital Social Networks Trust building Shared vision	SC SC_n SC_t Sc_v	KMS helps to expand and establish trustworthy personal networks, and enhance organizational vision, mission, and goals sharing. Social Capital comprises three components: network, trust, and shared vision.

**Table 4.6 Variables and Descriptions in This Study**

#### 4.4.5 An Initial Version of the Survey Questionnaire

Following the recommendations of Straub (1989) and Straub, Boudreau, & Gefen (2004), previously validated instruments were utilized wherever possible.

#### SCALES FOR PU, EOU, AND NORMS

The typical measurement scales for perceived usefulness (PU) and perceived ease of use (EOU) come from Davis (1986, 1989), and were adapted and simplified to two items for each construct in their research (Straub, Limayem et al. 1995). The titles of these two-item versions, in turn, were adapted for this research.

The scale for measuring “social norms” was taken from (Lucas and Spitler 1999), and given minimal adaptation in wording. The scales comprise four items, two items for each dimension respectively.

A seven-point Likert-type scale was designed for the scales “perceived usefulness”, “perceived ease of use”, “social norms”, where 1 stands for ‘strongly disagree’ and 7 for ‘strongly agree’.

It is worth noting that these well -established scales were subjected to further validation because they have undergone some degree of adaptation and modification and are to be used in new contexts.

#### **SCALES FOR PERFORMANCE-RELATED USE OF KMS**

The development of the scale items for measuring the performance-related use of KMS follows a procedure used in developing multidimensional measures for system-use constructs by Doll & Torkzadeh (1998). In the procedure, for each of the dimension of the multidimensional constructs, a set of typical activities is identified by reviewing the relevant literature: these activities are then transformed into corresponding questions one by one (Doll and Torkzadeh 1998).

In terms of item development for the two sub-constructs of the performance-related use of KMS, the identified activities are transformed into corresponding questions one by one by simply using the sentence model “I use KMS to do (activity)” (Doll and Torkzadeh 1998). In total, ten original questions for measuring the two sub-constructs were generated, four items for the information-related use of KMS and six items for the interaction-related use of KMS respectively. A five-point Likert-type scale was used to gauge the perceived extent of use of KMS in the selected activities where 1 stands for ‘Not at all’, and 5 for ‘A great deal’; “N/A” was used for the activities which were not available to the participants.

#### **SCALES FOR INFORMATION QUALITY**

There exists a variety of survey instruments (questionnaires) for measuring information quality in IS literature (e.g., Zmud and Boynton 1991, Garrity and Sanders 1998). Compared to the other information quality instruments, the survey questionnaire developed by Goodhue (1998) was considered to be closest to the basic requirements of measuring information quality in this study. This was because the questionnaire was

based on the task-technology fit theory, and was designed for evaluating information quality for knowledge work (e.g., non-routine decision making and problem solving). Moreover, the instrument has been validated carefully using a sample of knowledge workers (i.e., managers) (Goodhue 1998).

The original Goodhue survey instrument (Goodhue 1998) examined three aspects of information systems quality, i.e., information content, information services, and systems quality. The instrument comprised 24 items and eight dimensions, the right level of detail, accuracy, meaning, locatability, accessibility, assistance, ease of use, and systems reliability (Goodhue 1998). The scale, however, had to be adapted and modified so as to be applicable to this study. As a result, 13 validated items were retained, four questions measuring information content quality, seven questions measuring information services quality, and two questions measuring system quality respectively. A seven-point Likert-type measure was designed for the scales, where 1 stands for 'strongly disagree', and 7 for 'strongly agree'

#### **SCALES FOR SOCIAL CAPITAL DEVELOPMENT**

The social capital development scales measure the impact of KMS on three components of social capital construct, social networks, trust, and shared vision (Nahapiet and Ghoshal 1998; Tsai and Ghoshal 1998). The development of social capital scales in this research was based on the survey instrument of social capital developed by Tsai and Ghoshal (1998). However, while the Tsai and Ghoshal's (1998) instrument was used at an organizational unit level, the survey instrument in this study is required to measure social capital at individual level, and specifically emphasize the influence of KMS.

Measuring structural aspects of social capital at an individual level has been an important issue in previous research on social capital. A variety of measures have been developed, such as range of individuals' social networks (e.g., Meyerson 1994), network structure (Coleman 2000), and individual's position in social networks (Burt 1992; Burt 1997). For instance, Higgins and Nohria (1999) operationalized an individual's social capital as the number of ties with different national subsidiaries of the firm that the respondent reported. Flap and Boxman (1999) calculated managers' social capital based on the number of work contacts and the number of association

memberships of managers. In an empirical investigation of the impact of social contacts on managers' incomes, Meyerson (1994) measured an individual's social capital by the number of strong ties a subject possessed.

In this study, the impact of KMS on the user's social networks was measured by the change of the range of end-users' social connections within the organization which resulted from the use of KMS, i.e., measuring the degree to which a person believed that using KMS would expand his or her work connections, including knowing more people, expanding his or her information and knowledge networks and maintaining close relationships with friends (Cross, Parker et al. 2001; Cross, Borgatti et al. 2002). Four questions were developed for measuring the impact of KMS on individual social networks.

With respect to measuring the impact of KMS on the level of trust among the colleagues, it was felt important to measure the levels of trust emerging in new developed social connections through KMS. Where the range and connectivity of an individual's social networks was subjected to change due to the KMS use, the attributes of the new relationships should consequently be assessed. Three items for measuring the level of trust were generated by modifying the items developed by Tsai and Ghoshal (1998).

The scale for the impact of KMS use on the level of shared vision was developed based on a similar scale produced by Tsai and Ghoshal (1998). Changes were made to the original scale to meet the requirement of this study, including moving the scale from an organizational unit level to an individual level, and emphasizing the influence of KMS. As a result, two items were generated.

In sum, the original social capital scales comprised 9 items, in which four items for measuring KMS 's impact on individual's social network, three items for the impact on trust level, and two items for the impact on shared vision respectively. A seven-point Likert-type measure was designed for the scales, where 1 stands for 'strongly disagree' and 7 for 'strongly agree'.

#### **SCALE FOR FUNCTION-RELATED USE OF KMS**

The study uses self-reporting measures of current use for each of a set of selected functions provided by KMS. Lucas and Spitler (1999) note that

*“For a complex system with many possible functions, self-report measures may be the best alternative available.” (p. 298).*

The comment was considered to be applicable to KMS.

The extent to which the users use eight selected functions of KMS in their daily work was measured by a six-point scale, where 0 stands for ‘N/A’, 1 for ‘Extremely light’, and 5 for ‘Extremely heavy’; the choice of “N/A” was for those functions that were not available to the participant. The eight functions of KMS mentioned in the survey were email, video conferencing, knowledge repositories, information and knowledge distribution, location of/access to expertise and/or experts (i.e., Yellow pages), online discussion forums, online virtual communities and virtual collaboration and/or cooperation (see Section 2.5.3).

In addition, eleven items for collecting personal and organizational information and one open question for additional comments were included in the initial version of questionnaire. The complete initial version of questionnaire, which consists of 60 items, can be found in Appendix I.

In sum, this section was dedicated to the survey questionnaire development, which resulted in an initial version of questionnaire with sixty items. In the following section, a pilot study, also referred as a preliminary study, is undertaken to pre-test the initial questionnaire. The results of the preliminary study helped in simplifying, purifying, amending, and modifying the initial questionnaire.

## **4.5 Preliminary Study**

From a viewpoint of empirical study, a preliminary study is a necessary process to validate a newly developed or adapted survey questionnaire so as to ensure basic psychometric properties of the new questionnaire, and to lead to an instrument that could be formally validated (Straub 1989). A complete process of preliminary study



involves data collection with the initial questionnaire, and data analysis on the sample data. The software package SPSS 11 was used to conduct the data analysis (Coakes and Steed 1999).

#### **4.5.1 Data Collection**

To collect data to validate the initial version of the survey questionnaire for this study, a web-based survey was conducted via a professional Internet survey service provider—SurveyGold ([www.surveygold.com](http://www.surveygold.com)). The formatted electronic version of the questionnaire, with an introduction page and a statement of the ethics approval of the study, was hosted on a university server. An invitation letter with a link to the survey was then sent to the selected population via email. Professionals (e.g., academics, professional staff, and managerial staff), who work for high-tech. companies, financial institutions, educational and research institutions, government agencies, and professional services were assumed to be the suitable target population. These people are regarded as knowledge workers (Schultze 2003).

Just before the questionnaire, a description of the functionality of KMS was included in order to improve the validity of responses. The instruction page included a brief introduction to KMS and KMS usage and a list of selected KMS functions. Each potential participant was asked if he or she used any of the listed functions in his/her current organization. If answer was ‘yes’, the potential participant could continue the journey to complete the survey, otherwise, he or she could exit the survey.

A convenience sampling frame was constructed by means of available public Internet resources, such as listservs, professional discussion groups, and university academic staff directories (Simsek and Veiga 2001). 376 invitations were sent out, and 74 valid responses were obtained. The response rate was 19.68%, achieved with the help of a follow-up reminder email sent one week later. The response rate was considered normal in Internet survey (Simsek and Veiga 2001).

#### **4.5.2 Sample Characteristics**

54% of respondents were male, 46% were female. Respondents with postgraduate qualifications accounted for 67.57% of the sample (bachelor degree and above holders

accounted for 93%) suggesting that they were knowledge workers (Drucker 1999; Horibe 1999). 63.5% of the respondents came from the education and research sector, and the next two largest sectors for sample were government agencies (21.6%) and high technology companies (8.1%). 59.4% of the respondents classified their positions as academic, whereas professionals accounted for 27%, and managerial staff 10.8%. In addition, 89.2% of respondents had over one year experience in using KMS (56.8% respondents' KMS experience had over five years) and 75.7% of respondents came from organizations with over 500 employees.

The sample characteristics are summarized in Table 4.7.

<b>Sex (%)</b>	
Female	46%
Male	54%
<b>Highest Educational Level (%)</b>	
Secondary Qualification	1.35%
Associate	1.35%
Bachelor	25.68%
Postgraduates	67.57%
Other	4.05%
<b>Years with Current Organization (%)</b>	
Less than one year	9.46%
1 – 2 yrs	17.57%
3 – 5 yrs	22.97%
6 – 10 yrs	21.62%
More than 10 yrs	28.38%
<b>Position in Current Organization (%)</b>	
Academic	59.40%
Managerial	9.46%
Professional	27.03%
Administrative/Clerical	1.35%
Other	2.70%
<b>Industry (%)</b>	
High Technology/Computers/Telecommunication	8.11%
Banking/Finance/Insurance	1.35%
Education/Research	63.51%
Government Agencies	21.63%
Health and Social Services	1.35%
Professional Services (Legal, Accounting, Consulting)	1.35%
Other	2.70%
<b>Years of Experience in Using KMS within Current Organization (%)</b>	
Less than one year	10.81%
1 – 2 yrs	16.22%
3 – 5 yrs	16.22%
More than 5 yrs	56.76%
<b>Organization Size (number of employees, %)</b>	
Less than 500	24.32%
501 – 1,000	16.22%
1,001 – 5,000	50%
More than 5,000	9.46%

(Source: summarized from the outputs of SPSS 11)

**Table 4.7 Demographics of the Sample for Pilot Study**

### **4.5.3 Data Analysis**

First at all, a viable data analysis strategy was needed to guide the analytical process. According to Churchill (1979), exploratory factor analysis (EFA) and internal consistency assessment can serve the aims of preliminary study to purify and revise the initial items of measurement scales. In this preliminary study, data analysis followed a

similar procedure to that used by Doll and Torkzadeh (1998), but amendments were made to the procedure to include the examination of reliability before the EFA, in order to strength the preliminary study. As a result, the data analysis in this preliminary study followed the subsequent process of testing. Corrected-item total correlation analysis was used to purify the questionnaire construct by construct (Doll and Torkzadeh 1998), followed by an internal consistency assessment with Cronbach's alpha (Churchill 1979; Hinkin 1998). EFA was then used to identify items that were not factorially pure (Doll and Torkzadeh 1998).

Validity criteria were needed to judge the data analyses. There exist widely used validation heuristics for these analyses. For analysis with corrected-item total correlation, items should be dropped if their corrected-item total correlation is  $< 0.5$  (Doll and Torkzadeh 1998). For analysis with internal consistency assessment with Cronbach's alpha, as a rule of thumb, the acceptable Cronbach's alpha value can be as low as 0.5 or 0.60 for exploratory research, and 0.70 for confirmatory research (Churchill 1979; Straub, Boudreau et al. 2004). For EFA, items that loaded on more than one factor at 0.40 or above should be dropped (Doll and Torkzadeh 1998).

#### **4.5.3.1 Internal Consistency and Reliability**

Following the data analysis strategy for the preliminary study, corrected-item total correlation was used to purify scale items (Doll and Torkzadeh 1988; Doll and Torkzadeh 1998), followed by the calculation of coefficient alpha. According to the validation heuristics, items should be dropped if their corrected-item total correlation is  $< 0.5$  (Doll and Torkzadeh 1998), and coefficient alpha = 0.60 is regarded as an acceptable threshold value for the early stages of basic research (Churchill 1979).

Table 4.8 shows the analysis results generated by corrected-item total correlation analysis and the internal consistency assessment with Cronbach's alpha. The retained number of questions and corresponding Cronbach's alphas are in parenthesis next to the original values. Examination of the reliability coefficients indicates that the majority of the scales have an acceptable reliability (i.e.,  $\alpha \geq 0.60$ ) except for the scale for Perceived Usefulness ( $\alpha = 0.57$ ). Although reliability of this scale was marginally lower than the 0.60 threshold value, it was decided to keep the perceived usefulness

construct and its scale because it was a well-established and widely accepted construct and scale. The results, therefore, suggest that these survey instruments has shown acceptable internal consistency (Goodhue 1998). It is quite clear that the scale reliabilities were improved following the purification of items.

Scales	Number of Questions	Cronbach's Alpha
KMS-use	10(7)	0.88(0.89)
Info_Use	4(3)	0.78 (0.80)
CC_Use	6(4)	0.83(0.87)
INFOQ	11(10)	0.93(0.94)
Info_c	4(3)	0.80(0.87)
Info_s	7	0.95
SC	9(6)	0.89(0.86)
SC_n	4(2)	0.75(0.79)
SC_t	3(1)	0.82
SC_v	2	0.85
EOU	2	0.85
PU	2	0.57
NORMS	4	0.81
<b>Note:</b> The reliabilities and number of questions in parentheses indicate status after one or two questions were dropped according to the score of corrected item total correlation (<0.5). Refer to Appendix I for the actual wording of the questions kept and dropped.		
<b>KMS-Use:</b> Performance-related use of KMS	<b>Info_Use:</b> Information-related Use of KMS	<b>CC_Use:</b> Interaction-related Use of KMS
<b>INFOQ:</b> Information Quality	<b>Info_c:</b> Information Content Quality	<b>Info_s:</b> Information Services Quality
<b>SC:</b> Social Capital Development	<b>SC_n:</b> Social Networks	<b>SC_t:</b> Trust building
<b>SC_v:</b> Shared Vision	<b>EOU:</b> Perceived Ease of Use	<b>PU:</b> Perceived Usefulness
<b>NORMS:</b> Social Norms		

(Source: summarized from the outputs of SPSS 11)

**Table 4.8 Reliabilities of the Scales**

#### 4.5.3.2 Exploratory Factor Analysis

After initially purifying the measures, exploratory factor analysis (FA) was used to further simplify and purify the scales items, and analyze and confirm the factorial structure of constructs using SPSS 11 (Churchill 1979; Coakes and Steed 1999). Three scales, i.e., performance-related use of KMS (KMS-Use), information quality (INFOQ), and social capital development (SC), were subjected to EFA, specifying factors for each scale, and choosing Principal Components and oblique rotation (Oblimin with Kaiser Normalization) because the underlying factors for each construct might be correlated with each other (Tabachnick and Fidell 1996).

The interpretation of factors and factorial structure for each construct can be achieved through item loadings from the loading matrix output from SPSS 11. The item loading is the correlation coefficient between an item and its factor (for orthogonal rotation), or a measure of the unique relationship between the item and its factor (for oblique rotation) because the factors correlate in the later case.

Table 4.9 shows the loadings of remaining items (the items with loadings less than 0.40 or cross loadings greater than .40 were dropped) on the assigned factors. Examination of the loadings matrix indicate that there exists a clear and interpretable factor structure within each given construct, which is in accordance with the assumed factor structure for each construct, and the loadings of all the items on their respective assigned factor can be considered excellent (in excess of 0.71) (Tabachnick and Fidell 1996), meaning that they were pure measures of the respective factors. In addition, while the sample size was not big enough (N=74), the system outputs of Kaiser-Meyer-Olkin measure of sampling adequacy and the Measures of sampling adequacy (MSA) were well above the acceptable levels (i.e., .6 and .5 respectively) for all the three constructs, which ensured that the results of factor analysis could be considered acceptable (Coakes and Steed 1999).

Item code	Info_use	CC_use	Info_c	Info_s	SC_n	Sc_t	SC_v
KMS- Use Q12 Q13 Q14 Q18 Q19 Q20 Q21	0.81 0.94 0.73	0.74 0.88 0.90 0.85					
SC Q23 Q24 Q26 Q27 Q29 Q30					0.82 0.90	0.95 0.80	0.75 0.94
INFOQ Q39 Q41 Q42 Q43 Q44 Q45 Q46 Q47 Q48 Q49			0.86 0.82 0.95	0.75 0.93 0.94 0.95 0.86 0.85 0.80			
KMS-Use: Performance-related use of KMS		Info_Use: Information-related Use of KMS			CC_Use: Interaction-related Use of KMS		
INFOQ: Information Quality		Info_c: Information Content Quality			Info_s: Information Services Quality		
SC: Social Capital Development		SC_n: Social Networks			SC_t: Trust building		
NORMS: Social Norms		SC_v: Shared Vision Development					

Extraction Method: Principal Component Analysis. Rotation Method: Oblimin with Kaiser Normalization (Source: summarized from the outputs of SPSS 11)

**Table 4.9 Rotated Factor Loadings for KMS-Use, SC, and INFOQ**

#### 4.5.4 Revising the Survey Questionnaire

The preliminary study focused on the pre-test of the initial survey questionnaire, and the results have shown acceptable psychometric properties of the instrument, offering some promise for further study. The preliminary study provides empirical evidence for endorsing the survey instrument and helps its evolution.

During the preliminary study, the items of the questionnaire were purified and reduced as the results of statistical analyses, such as corrected-item total correlation analysis,

reliability analysis, and factor analysis. The preliminary study resulted in the retention of seven items for the two components of the performance-related use of KMS (KMS-use), ten items for the two components of Information quality (INFOQ), six items for the three components of Social capital development (SC) construct, two items for Perceived usefulness (PU), two items for Perceived ease of use (EOU), and four items for Social Norms (NORMS).

Based on experience from the preliminary study and based on the literature, the questionnaire items were reviewed carefully resulting in some modifications to items in the performance-related use of KMS (KMS-use) and social capital development (SC) scales; the other scales remained unchanged.

For the performance-related use of KMS (KMS-use) scale, Q12, Q13, Q19, Q20, and Q21 were retained, Q18 was dropped because this question was considered to be very similar to that of Q17, which was dropped during the preliminary study. The wording of Q14 has been amended slightly as to make it more accurate. In addition, a new sentence was added to the scale for measuring user behaviour in using KMS to seek work-related information and advice, considered to be important for knowledge work. As a result, the scale for measuring KMS-use was revised, whilst remaining at seven items. The revised scale can be found in Table 4.12.

With regard to the scale for social capital development, substantial modification was made to the scale based on the literature. While the two items for shared vision remained unchanged, the scale for personal social networks and the scale for trust building were reformulated, for two reasons. Firstly, the focus of an individual's social networks was extended to include work-related connections, such as an advice-network, because information and advice-networks play a more obvious and significant role in knowledge creation and sharing (Cross, Parker et al. 2001; Cross, Borgatti et al. 2002). Secondly, it was felt that the scale development needed to be orientated towards a professional work relationship, that is the knowledge work environment rather than the general work environment. As a result, questions Q22, Q23, and Q24 have been combined in order to capture the impact of KMS on an individual's social connections within their organization. Furthermore, a new statement, "The KMS helps expand my information & advice-networks within the organization, from which I can regularly seek



work-related information and advice” was added to emphasize the development of work-related networks.

As mentioned earlier, trust is critical to knowledge management and knowledge work, and is viewed as the bandwidth of information and knowledge networks (Sveiby 1997; Cohen and Prusak 2001). In recent research on trust building in professional work relationships, Lewicki and Bunker (1996) identified three types of trust, namely calculus-based trust, knowledge-based trust, and identification-based trust. Unlike calculus-based trust which is grounded in the fear of punishment or in the rewards, and identification-based trust which relies on a much closer relationships between people, knowledge-based trust builds on the information, knowledge, and understanding of others, which can develop with interactions with fellow workers. Lewicki and Bunker (1996) note that knowledge-based trust relationships are characterized as being more productive than other relationships, and “*relationships at work are often knowledge-based trust relationships*” (p.125). The basic activities that build and strengthen knowledge-based trust involve information-related behaviour, such as information access, sharing, and dissemination, and mutual understanding-enabled behaviour, such as communication, collaboration, joint projects and shared experience (Lewicki and Bunker 1996). As KMS provides a platform for people to seek information, share experiences, maintain communication and interaction with each other, enabling mutual understanding between colleagues, it is reasonable to believe that the KMS can play a critical role in enhancing knowledge-based trust development in professional working relationships.

For this study, it was felt necessary to shift focus to knowledge-based trust, and to formulate a new trust measure based on those activities that may build and strengthen knowledge-based trust, to replace the scale adopted from Tsai and Ghoshal (1998).

The revised scales for measuring the performance-related use of KMS and the potential impact of KMS on social capital development (social networks, trust building, and shared vision) in knowledge work environment are shown in Table 4.10. The scales were subjected to a further validation.

In order to assure the content validity of the revised survey questionnaire, all the items were reviewed one by one with several university academics, including my supervisor

and other university staff, against the corresponding definitions of the constructs (Schriesheim, Powers et al. 1993; Hinkin 1998). As the result, several wording problems were identified and corrected.

<div>KMS-Use Info_Use</div> <div>CC_Use</div>	<div>I use the KMS to access online documents /knowledge repositories. I use the KMS to distribute information. I use the KMS to search for people who might have the information/expertise I need. I use the KMS to seek work-related information and advice.</div> <div>I use the KMS to maintain communication with colleagues. I use the KMS to discuss ideas, and/or exchange views/experience with colleagues. I use the KMS to cooperate with colleagues. I use the KMS to collaborate with colleagues.</div>	
<div>SC SC_n</div> <div>SC_t</div> <div>SC_v</div>	<div>The KMS helps me expand connections with people (i.e., make more acquaintances and friends) in my organization. The KMS helps expand my advice-network within the organization, from whom I can regularly seek work-related information and advice.</div> <div>The KMS enhances my understanding of colleagues' knowledge and skills. The KMS improves experience sharing between colleagues and me. The KMS enhances the mutual understanding between colleagues and me (i.e., through online information, communication &amp; interactions). The KMS helps me maintain regular communications with colleagues. Because of the information and experience I get through the KMS, I feel that I can rely on those from whom I seek work-related information and advice. Overall, the KMS improves trust level between colleagues and me.</div> <div>The KMS helps me and colleagues share same ambitions and vision. The KMS helps me and colleagues keep pursuing the collective goals and mission of the whole organization.</div>	
<div>KMS-Use: Performance-related use of KMS</div> <div>SC: Social Capital Development</div> <div>SC_v: Shared Vision</div>	<div>Info_Use: Information-related Use of KMS</div> <div>SC_n: Social Networks</div>	<div>CC_use: Interaction-related Use of KMS</div> <div>SC_t: Trust Building</div>

**Table 4.10 Revised Measures for KMS-Use and Social Capital Development**

## 4.6 The Survey

The initially validated survey questionnaire was subjected to further validation, and the developed hypotheses tested by a formal survey using the revised survey questionnaire. The second sampling, lasted for about eleven weeks, took place six weeks later after the first data collection. This second data collection was conducted via the same

professional Internet survey service provider—SurveyGold ([www.surveygold.com](http://www.surveygold.com)). The first electronic version of the questionnaire was replaced by the revised version, with an updated introduction page and a statement of ethics approval of the study by the University. The questionnaire was hosted on a university server, and an updated invitation letter with a link to the questionnaire was then sent to the selected population by email. Based on the experience from the pilot study, it was estimated that the questionnaire would take around 15 minutes to complete. The same target population, professional or knowledge workers, was chosen for the data collection.

The complete survey package, which comprises, an invitation letter, a reminder letter, an instruction page, and questionnaire, can be found in Appendix II.

As in the first survey, a description of the functionality of KMS was included in the instruction page of the survey to improve the validity of responses. The instruction page included a brief introduction of KMS and usage and a list of selected KMS functions. Each potential participant was asked if he or she used any of the listed functions in his/her current organization. If answer was 'yes', the potential participant could continue to complete the survey, otherwise the respondent exited the survey.

Whilst a convenience sampling frame was constructed by means of available public Internet resources, such as listservs, professional discussion groups, and university academic staff directories (Simsek and Veiga 2001), efforts were made to expand the target population to include more professionals from as wide a range of professional groups as possible, including computing and telecommunications, the pharmaceutical industry, the scientific community, librarians and government agencies, instead of focusing on university teachers as in the trial data collection. The respondents in the first round of data collection were omitted from this survey. One week after the invitation letter was sent, a follow-up reminder email was sent to encourage more responses, as had been done in the trial data collection.

In total, 2530 invitations were sent out during a period of around eleven weeks. Of these 155 invitations was rejected because of wrong or abandoned email addresses. 362 usable responses were received. After deduction of the dead mails, the response rate is 15.24%. While it is a reasonable response rate for an internet-based survey (Simsek and

Veiga 2001), there were several reasons for the lower response rate, according to the responses received. Firstly, most people approached were professionals with tight time schedules. Secondly, some organizations banned any employee from participating in surveys during their working time; several respondents informed me that they could only complete my survey during the weekend. Thirdly, companies concerned about Internet security rejected any survey invitations or deleted any unexpected emails with attachments. There was no missing data in the sample.

## **4.7 Chapter 4 Summary – Research Methodology**

This chapter started with a discussion of the major research approaches and philosophies in IT/IS study. A variety of research approaches were found in IS research, with three fundamental research perspectives, namely the positivist, interpretive, and critical perspectives. Of the three, the positivist was discussed in more detail because the positivist approach is recognized as the major research perspective in information systems research. The positivist perspective was chosen as being suitable for this study. In terms of research methods and techniques, surveys, particularly Internet surveys, have been chosen to be the major data collection tools for this study, and discussed in detail. Structural Equation Modelling (SEM) techniques have been addressed in detail as the second-generation statistic techniques which have been chosen as the major statistical analysis tools for this study.

A complete development process of the survey questionnaire (scale development and validation) has been discussed and articulated in this chapter. Two trials of data collection, targeting end-users of KMS from various organizations in a variety of countries, have been conducted by means of Web-based survey, resulting in good quality results in terms of the size of the sample, the quality of the respondents, and the diversity of organizations and countries. A pilot study of the newly adapted and developed survey questionnaire has been conducted, resulting in revised a survey questionnaire with reasonable psychometric properties, which is considered to be ready for further formal construct validation. Using the revised survey questionnaire, a larger scale of survey has been conducted, resulting in a quality sample for further study.

The next chapter is dedicated to data analysis, validation the survey instrument and testing of the hypotheses.