

Bridging the Unspannable Chasm: Qualitative Knowledge Construction for Engineering Systems

Jason E. Bartolomei¹
United States Air Force, Dahlgren, VA, 22448

Susan S. Silbey² Daniel E. Hastings³ Richard de Neufville⁴ Donna H. Rhodes⁵
Massachusetts Institute of Technology, Cambridge, MA, 02

Copyright © 2009 by Jason E. Bartolomei. Published and used by MIT ESD and CESUN with permission.

The understanding of engineering systems demands diverse knowledge. Managers must consider the social and technical, the observable and unobservable, the concrete and abstract, and the animate and inanimate to successfully describe and purposively shape the system. Therefore, the complexity of engineering systems requires hybrid approaches for constructing our knowledge of such systems. This research attempts to bridge what has frequently been described and experienced as an unspannable chasm between the methods of social science and engineering science.

Key Words: engineering systems, social science research methods, modeling, grounded theory, quantitative analysis, design structure matrix, systems engineering

1. Introduction

Too often overlooked in the chaos of disciplines (Abbott 2001, Abbott 1988), there are important similarities between social science and engineering science. Both share an interest in the structure of a system, the relationships between a system and its environment, and system behavior. Nonetheless, a perceived vagueness of social science by engineers prevents them from bridging epistemological, disciplinary, and departmental boundaries. Engineers use principles of physical science as a basis for description, analysis, and decision-making. Physical science has produced

¹ Research Affiliate, MIT Engineering Systems Division, United States Air Force, 520 Caffee Circle, Dahlgren VA 22448

² Head, Anthropology, Leon and Ann Goldberg Professor of Sociology and Anthropology, MIT, 16-233, 77 Massachusetts Ave., Cambridge MA 02139

³ Dean, Undergraduate Education, Professor of Engineering Systems and Aeronautics and Astronautics, MIT, 4-110, 77 Massachusetts Ave., Cambridge MA 02139

⁴ Professor, Civil and Environmental Engineering and Engineering Systems, MIT, E40-245, 77 Massachusetts Ave., Cambridge MA 02139

⁵ Senior Lecturer, Engineering Systems Division, Principal Research Scientist, SEArI and LAI, MIT, E38-572, 77 Massachusetts Ave., Cambridge MA 02139

reliable laws and rules, usually mathematically supported models and equations that purportedly model nature and consequently enable design of new processes and objects. The theoretical grounding of this knowledge results from hundreds of years of modern empirical physical science and has become so reliable and formulaic that engineers can routinely use this knowledge to build and analyze complex physical systems. However, when it comes to the social and organizational aspects of systems, social science is still in its infancy. It is only 150 years old and still lacks equally robust, general statements, rules or laws. Although theoretical synthesis of social scientific knowledge is still relatively incomplete, an insufficient foundation for strong predictive models (Rein 1999), social scientists have developed several broadly applied and reliable techniques for describing and analyzing social action and culture.

Nonetheless, neither engineering nor social science has satisfactorily addressed the behavior of the social aspects of systems intersecting with the technical aspects. Engineers pay inadequate attention to the human components of systems. Social scientists too often fail to account for the role of technological and physical constraints and opportunities although a growing literature in science studies is filling this gap (e.g. Latour). By adopting methods of ethnographic fieldwork and grounded theory construction, two well-respected modes of social science research to explore both the social and technical components of systems, this research hopes to make transparent what has too often been the black box of systems analysis.

Qualitative Knowledge Construction describes a mode of incorporating qualitative information about an engineering system. Information can be collected through diverse modes: interview, observation, and documentation. The information is converted into textual form that is subsequently coded by the categories (nodes, relationships, attributes). Once coded, the data is automatically inserted into a rich data structure. In this manner, the system is visually represented by the distribution of the values of the cells; the evidentiary data supporting that representation is contained within the cells. By developing a program for translating textual reports of observations and interviews into coded data capable of being systematically and quantitatively analyzed, the researcher can translate qualitative information into quantifiable data. Not only can social scientific methods improve the descriptive validity of a complex system, but they can be used to improve an engineer's prescriptive ends as well.

In practical terms, QKC offers a means for better book keeping when modeling an engineering system by providing traceability of assumptions to data sources. Dong (2001) found that that much of the information surrounding engineering system resides in the minds of the humans involved in the system, not in technical documentation. QKC provides a means of capturing this knowledge and thus providing a transparency of assumptions and traceability to sources that rarely exists in the study or practice of engineering. By explicitly establishing this transparency and traceability, QKC removes epistemological barriers for the scientific study of engineering systems thus leading to better theory building, design, and management.

2. Qualitative vs. Quantitative Methods for Constructing Knowledge

There are many differences between qualitative and quantitative research methods. For simplicity, in quantitative social science, statistical methods are used to draw conclusions about a particular social phenomenon from analyses of variation, correlation, and other forms of association within large data sets composed of quantitatively represented information. Qualitative social science

refers to methods of “conducting inquiry that are aimed at discerning how human beings understand, experience, interpret, and produce the social world” (Sandelowski 2004). One important difference between quantitative social science and qualitative social science is the amount of information that the researcher ostensibly knows, or thinks he or she knows, ahead of time, before the information is acquired. Quantitative studies hypothesize and model what is expected, developing modes of data collection through fixed categories that purport to measure the variation in predetermined and quantified variables. Becker (2001) explains that, for qualitative researchers, there are often surprises in the data acquired, and much of the data collection processes (in the field) and forms of analysis (of the textual representations) are designed to be open to unanticipated information and connections. The qualitative approach contrasts with quantitative researcher’s a priori “knowledge” of the data of interest; the researcher engaged in qualitative data collection is not limited to the questions on a survey as additional data and unexpected insights emerge through work in the field.

For a qualitative researcher, the process of collecting data involves the mutual construction of data between the researcher in concert with social actors within the system. In doing so, the researcher sets out opportunities for sharing rich detailed descriptions or observing actors at work. Rich detailed data or “thick” data are written descriptions of events observed by researchers, extensive accounts of personal experience from respondents, and records that provide narratives of experience (usually transcriptions of interviews). In addition to participant observers’ field notes, other accounts may produce rich detailed data. Rich data includes thoughts, feelings, actions, and context (both material and relational). Thick, layered descriptive information enables the analyst to trace events, to delineate steps and stages in a process, and to make comparisons. From the thickly described and detailed information, we can begin to construct our model of the engineering system. Many engineers and managers are familiar with case studies and case histories, as these are often the subjects of engineering and business literature. They may, however, be less familiar with ethnographic methods of data collection and grounded theory modes of data analysis.

3. Grounded Theory

When grounded theory originated as a social science method, researchers wanted to develop a more systematic approach to analyze the wealth of qualitative data that was being collected through observation and interviews. Glaser and Strauss (1967) offered grounded theory as a credible methodological basis for theory building using qualitative data. The aim of ground theory methods is to generate theory from the ground up. This is done by creating abstract concepts and postulating relationships through inductive examination of empirical data. Rather than postulate relationships as in a hypothesis and then operationalize the concept into measurable units that are then tested with empirical observation, grounded theory attacks a collection of textualized data as an open field of possibilities. Grounded theory is described as a “flexible, yet systematic mode of inquiry”, “directed, but open-ended”, and an enabler of “imaginative theorizing”.

In the simplest terms, grounded theory is primarily, though not exclusively, inductive. Rather than beginning with a model and observing empirical phenomena to determine whether they align with a hypothesized relationship, grounded theory steers clear of this deductive approach by starting with observations from which categories of similarity and difference are developed and then aggregated into a model or hypothesis of a phenomenon.

4. Bridging the social and technical, qualitative and quantitative

The advantage of qualitative social science methods like grounded theory is an emphasis on data collection, documentation, transparency of assumptions, and systematizing of data analysis. A methodology for systematically capturing this knowledge by transcribing interviews and systematically analyzing these data has several advantages over current systems engineering methods for model building.

For classically trained engineers and managers, the fuzziness of theory building by observing social interactions must seem a bit too intangible and intractable. Nonetheless, systems engineers recognize that engineering is a social process and that much of the knowledge concerning the design process and the technical artifact is in the minds of the people creating and enacting the system. Thus, some specific methods may be a useful way for constructing systems-level models. The methodology proposed here is called **qualitative knowledge construction (QKC)**.

5. Procedures for Qualitative Knowledge Construction

Like grounded theory, QKC offers an iterative, systematic process for researching a complex system that consists of a series of steps that include the following:

- Identify a system of interest
- Define objectives for analysis
- Collect data
- Code raw data
- Organize coded data into a systems model
- Examine model for missing and/or conflicted data
- Resolve missing and/or conflicted data
- Perform analysis
- Iterate

The details for each step are described in the sections below. The paper concludes with an application of QKC on a specific example.

Identify systems of interest:

The determination of system type is the first step in the methodology. For engineering systems, Bartolomei 2007 presents an ontological framework called the Engineering Systems Matrix (ESM) that defines classes of nodes and types of relations useful for modeling an engineering system.

In addition to identifying the system type, researchers should develop a tentative formulation of various characteristics of the system, namely defining the system boundaries, from what perspectives the system is modeled, developing strategies for observing the system of interest, and where can data about the system be found. As data are collected and analyzed, the details of these assumptions will be iteratively refined and improved.

QKC can be used to study other types of systems. For example, a systems biologist might want to construct a systems-level model of a physiological system. In this case, the only difference

procedurally is the researcher would use a conceptualization that describes biological systems in order to classify and relate the system components.

Define objectives for analysis: For any model, it is important to determine the objectives. In analyzing large-scale systems (e.g. the F/A-22 product development system), the modeling objectives might not require a systems model representing several levels of decomposition of the thousands of employees and the millions of technical components. Rather, the system modeler may be interested in questions that can be understood by abstracting the complexity of the system to simplify the modeling process. Because QKC is an iterative method, details can be added as the interests of the researcher change over time.

Collect data: The process of data collection for engineering system involves a variety of data types and methods of data gathering. Because much of the knowledge about a complex system resides in the minds of the human agency involved or surrounding the system, qualitative social science methods for eliciting data through interviews are central to the methodology. As such, researchers must identify subjects knowledgeable about the system, interview subjects using open-ended questions, and transcribe interviews into text. It is important to note, that QKC is not limited to interview transcription as the only source of data. QKC can be applied to other types of data. These might include technical data used for computational models, engineering drawings, and systems documentation as well as program documentation, presentations, or other information pertaining to the system. Because QKC is a form of exploratory research, new sources of data will emerge through interviewing participants and observation of the system. As data is collected about a system, the research can begin the process of qualitative coding.

Code the Data: Qualitative coding in QKC is slightly different from what is described in grounded theory. In QKC, the process of coding begins by developing a coding classification a priori that is based on the system type. For systems classified as engineering systems, six coding classes are defined by the ESM modeling framework (Stakeholders, Objectives, Functions, Objects, Activities, and System Drivers). These classes serve as the basis for organizing the codes that emerge from the data. In QKC, codes take the form of system components (node) and the relationship between components (relations). The attributes of nodes and relations can be coded as well. In the spirit of grounded theory, researchers are encouraged to identify and record codes that are not easily classified in the ontology. These “orphan” codes can be later integrated into the system model or used as the basis for further examination using grounded theory.

Figure 1 illustrates these ideas. Take, for example, the construction of a systems-level model of a hospital. An analyst would collect data describing the system. This data might include transcripts of interviews with system relevant actors, systems models used by the hospital to manage processes, documentation of hospital protocols and standards, architecture drawings of the hospital infrastructure, organizational charts, email messages, photographs, and any other type of data describing the system. The data can then be coded (through conceptual or 'line by line' and then conceptual coding) and organized into the systems-level framework as illustrated in Figure 1. The figures on the left represent various types of data sources that describe the system of interest. The ones on the right represent codes derived from the data that will be organized into a systems-level model of the system. The different colors represent the class of code (green are stakeholders, blue are objects, etc) the small ovals on the far right represent codes with attributes. For example, a

stakeholder “John” may be defined in an interview transcript. From the transcript, it is learned that John is 6 feet tall. The attribute for storing John’s height can be defined and represented.

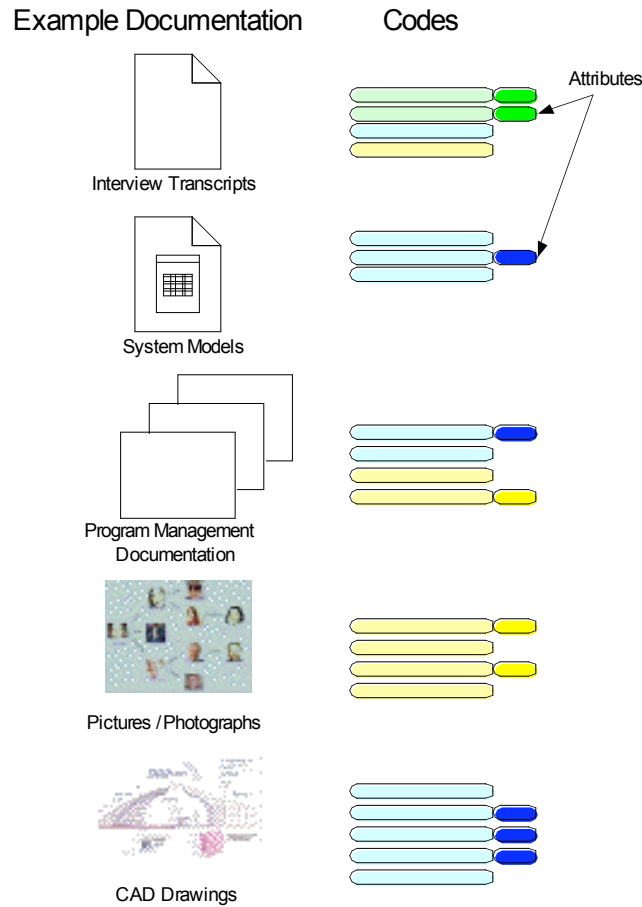


Figure 1: Coding in QKC

Figure 2 demonstrates the coding process. Take, for example, an excerpt from an interview with an actor called “Mary”. Mary is a manager of an engineering project. As a manager, Mary is classified as an external stakeholder within the system. On the left is a portion of a transcript of an interview with Mary. She is asked about the primary customer for her project. In the interview, Mary identifies “John” as one of the stakeholders in the system. On the right of the highlighted text is a code, “Stakeholder.John”, which identifies John as a new element in the systems model. In the same manner, other codes that emerge from the document are organized in the matrix as well. In this example, other codes include relations between John and Mary (John → interacts with → Mary) and Mary and John (Mary → interacts with → John). The code Stakeholder.John will be used each time John is mentioned in the data that the researcher collects. The figure shows other codes that emerge from the interview transcript as well.

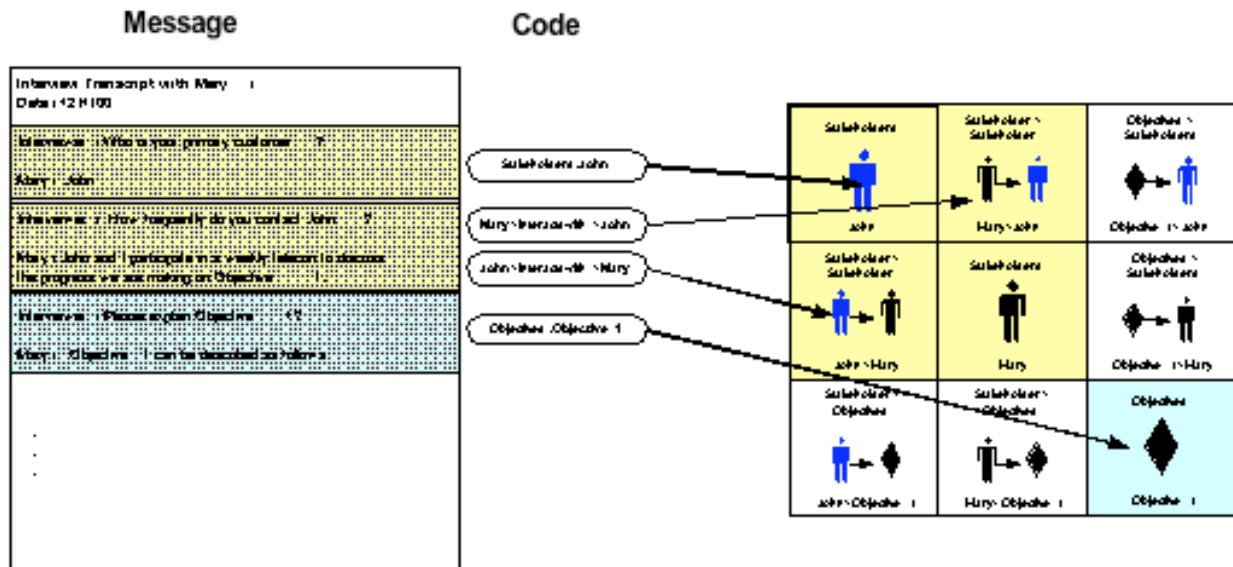


Figure 2: Example of Qualitative Knowledge Construction

Organize coded data in a systems-level modeling framework: After coding the various forms of data, the codes are to be organized into the systems-level framework that describes the system. From the example shown in the Figure 2, the code prefix “Stakeholder” signifies that John is a stakeholder and is represented in the corresponding upper left cell shown in the matrix on the far right. Similarly, the codes for relations can be organized in the matrix. This can be done by hand, through the use of computer spreadsheet software, or a customized database. A team of MIT researchers created a customized software application to streamline the QKC process for coding and creating a systems-level model.

Examine the model for missing/conflicting data. Once a systems-level model is constructed, the data can be examined to identify missing or incorrect information. Because each element of the model can be referenced to raw data (interviews, documentation, etc) researchers can invite others to review the data to verify the assumptions.

In the case of engineering systems, there is a foundational assumption that all elements within the system either contribute directly (or through other components) to the system goals. Therefore, any discontinuities in the data should be resolved. An example of a discontinuity is the identification of objects in the system that are not traced to functional components need to be reconciled. In this example, researchers might ask questions like, “have we missed any functions?” or “are these objects constituent components of the system, or not?” There are various graph theoretic search methodologies for identifying these types of gaps in the matrix.

Resolve missing data: Once missing or conflicting data in the model is identified, analysts must take action to resolve the conflicts. This is done through additional interviews, reviewing the raw data, and other similar actions.

Perform Analysis: Once a systems model has been developed to contain the qualitative data into quantifiable matrices, the researcher can apply various quantitative analytical methods for examining the system structure and behavior.

Iterate: Like grounded theory, QKC is an iterative process and researchers may be likely to perform several iterations of the methodology in the analysis of a complex system.

The next section presents an illustrative example that demonstrates the QKC methodology in modeling a generic, multistage supply chain.

6. Modeling a Supply Chain Using Qualitative Knowledge Construction: An illustrative example

In management science, the Beer Game has become a well-known tool for demonstrating the counter intuitive behavior of supply chains and the importance of information. The beer game is a simplified model of a basic supply chain that models the production, distribution, and delivery of beer. The model is an example designed to illustrate a variety of management principles ranging from the importance of information, human behavior, etc. This example demonstrates how a system analyst can take a textual description of the beer game and build an ESM using qualitative knowledge construction methodology.

6.A BEER GAME BASICS:

The Beer Game is a highly abstracted, simplified model of a multi-stage distribution system or supply chain. The system described in the Beer Game can be represented as an engineering system using the ESM framework (see Bartolomei 2007). The system consists of five stakeholders (Customer, Retailer, Wholesaler, Distributor, and Factory) each with a well-defined objectives (to minimize cost) that is calculated based on beer deliveries and inventory. The goal of the game is to simulate the dynamic behavior of supply-chain and highlight common challenges for supply chain management. Figure 4 illustrates the Beer Game layout mapping the supply chain from beginning to end.

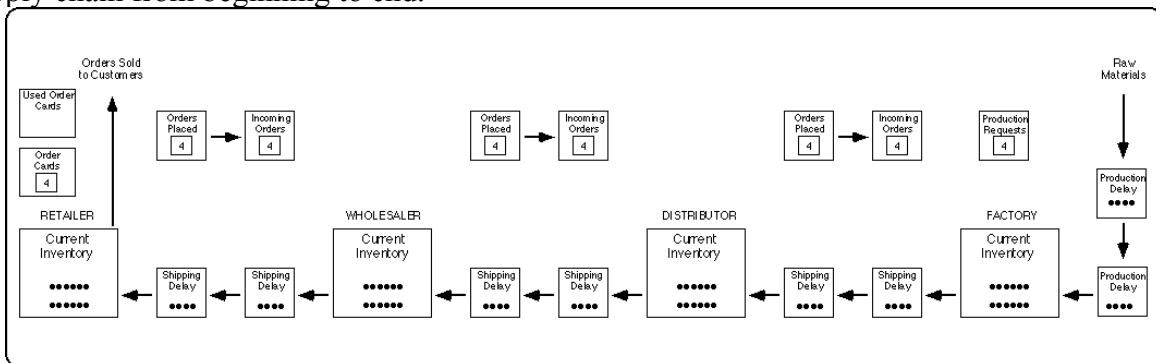


Figure 3: Beer Game Layout (Source Sterman 1994)

6.B APPLYING THE QKC METHODOLOGY:

Step 1. Identify a system of interest: The system of interest is a hypothetical supply chain for beer distribution defined by Sterman (1992). The system is an example of an engineering system as defined previously by Bartolomei (2007). The boundary of the system includes all components

that can be controlled by the Retailer, Wholesaler, Distributor, and Factory. All other components are considered exogenous to the system.

Step 2. Define analysis objectives: The modeling objective is to simulate the dynamics of the supply chain. This includes inventory dynamics, beer deliveries, and profits and losses.

Step 3. Collect Data: Numerous papers have been written to describe the beer game and many variants now exist. For the purpose of this example, the description of the system described in Sterman (1992) is used.

Step 4. Code Data: Using Sterman’s description of the Beer Game, the document is coded line-by-line using the QKC coding approach. An example of line-by-line analysis of the Beer Game text is shown in the figure below. The codes were generated using the comment feature of Microsoft Word.

Managers in an executive workshop playing the Beer Game at MIT.

Playing the Game

The game is played on a board that portrays the production and distribution of beer (figures 1-2). Each team consists of four sectors: **Retailer, Wholesaler, Distributor, and Factory** (R, W, D, F) arranged in a linear distribution chain. One or two people manage each sector. Pennies stand for cases of beer. A deck of cards represents customer demand. Each simulated week, customers purchase from the retailer, who ships the beer requested out of inventory. The retailer in turn orders from the wholesaler, who ships the beer requested out of their own inventory. Likewise the wholesaler orders and receives beer from the distributor, who in turn orders and receives beer from the factory, where the beer is brewed. At each stage there are shipping delays and order processing delays. The players' objective is to minimize total team costs. Inventory holding costs are \$.50/case/week. Backlog costs are \$1.00/case/week, to capture both the lost revenue and the ill will a stockout causes among customers. Costs are assessed at each link of the distribution chain.

The game can be played with anywhere from four to hundreds of people. Each person is asked to bet \$1, with the pot going to the team with the lowest total costs, winner take all. The game is initialized in equilibrium. Each inventory contains 12 cases and initial throughput is four cases per week. In the first few weeks of the game the players learn the mechanics of filling orders, recording inventory, etc. During this time customer demand remains constant at four cases per week, and each player is directed to order four cases, maintaining the equilibrium. Beginning with week four the players are allowed to order

- Comment [jb1]: Stakeholder: Retailer
- Comment [jb2]: Stakeholder: Wholesaler
- Comment [jb3]: Stakeholder: Distributor
- Comment [jb4]: Stakeholder: Factory
- Comment [jb5]: Stakeholders: Customers
- Comment [jb6]: Stakeholder: Customers >communicates to> Stakeholder: Retailer
Stakeholder: Retailer> delivers beer to> Stakeholder: Customer
- Stakeholder: Retailer>delivers product to> Stakeholder: Customer
- Comment [jb7]: Objects: Retailer:Inventory
- Comment [jb8]: Objects: Retailer:Inventory: Holding Costs,[t=0,\$.50]

Figure 4: Beer Game Coding Example

Examples of the codes are shown in the margin on the right. They include nodes (e.g. Stakeholder: Retailer), relations (Stakeholder: Retailer> delivers product to> Stakeholder: Customer) and attributes (Objects: Retailer Inventory: Holding Cost at time 0, \$.50 per unit).

Step 5. Organize coded data in systems-level modeling framework: Each of the codes and attributes can be represented in a model of the system using the ESM framework. Figure 6 is a matrix representation of the coded text organized using the ESM framework.

The northwest corner of the matrix shows the social interactions between the stakeholders in the game. These include the customer, retailer, wholesaler, distributor, and factory. Moving

analysis, discrete event simulation, agent-based simulation, or explore combinations of analysis. For example, the structure of the social network described in the text constrains how information passes between players. Using QKC researchers could explore alternate social network structures and information flows to see how they might affect the outcomes of the game.

QKC could also be used to explore game theoretic analyses of the beer game. Information could be collected about the attributes of the players of the game and their payoffs for success. This information could be used to explore alternate cooperative and non-cooperative scenarios.

Lastly, the rich data set produced through various iterations of the QKC methods could be used as a basis for model mixing. For example, ongoing research is exploring methods to combine system dynamics modeling with game theory model to explore how the dynamics of the beer game (via the system dynamics model) effect stakeholder interactions (via the game theory model) or vice versa.

6.C REAL WORLD APPLICATIONS OF QKC

Bartolomei (2007) presents a real-world example for how QKC can be used for studying product development systems for a Miniature Uninhabited Air Vehicle (MAV-PD). Qualitatively, the methodology led to a number of observations about the MAV-PD that serve as a basis for developing researchable questions for future research. Quantitatively, QKC provided a means to apply various well-established analysis methods such as classic Design Structure Matrix (DSM), network models, and a variety of other analytical methods for descriptive and prescriptive ends.

7. Comparing QKC and Grounded Theory

QKC contrasts with canonical grounded theory in several important ways. First, QKC begins with a well-defined preconception of the system qua system, based on disciplinary knowledge. The Engineering Systems Matrix serves as basis for researching a complex system by providing a classification framework for organizing knowledge of a system by class of objects (in the case of engineering systems: System Drivers, Stakeholders, Objective, Functions, Objects, and Activities) and type of relations (Signal, Material, Information, etc.). This is the first iteration of the classes of information used to describe the system. Thus, an analyst can collect data (in the form of interviews, photographs, Computer-Aided Design (CAD) drawing, system models, etc) that describe the system and by coding the data construct a systems-level model.

Second, QKC differs from inductive grounded theory in that it specifies variation in codes a priori. Some codes refer to the components of the system, e.g. stakeholders, objects, functions. Others refer to attributes or information about these elements of the system. In grounded theory, traditionally, these distinctions are relevant only if they emerge inductively.

Third, QKC differs from canonical grounded theory by locating unanticipated information, i.e. information that is beyond the originally stipulated objects, persons, and relationships, as orphan codes. Orphan codes can refer to information that cannot be representing in the existing modeling framework. For engineering systems, orphans might include information about abstract concepts concerning power of individuals, affect/emotion, or frustrations. By marking this unanticipated information as orphans, researchers are able to analyze how the emergent description varies from the hypothesized system. This allows for iterative improvement in modeling techniques over time.

In addition, by systematically incorporating orphan codes within original models future conceptualizations for complex systems will improve.

8. Conclusion

In summary, this paper presents a new procedure for constructing a systems-level model of an engineering system. The methodology is an improvement on existing systems approaches by explicitly using established qualitative methods to construct a systems-level model of the system. The process takes both qualitative and quantitative information surrounding the system and transforms this information into a quantifiable data structure that serves as a tool for better theory building, design, and management of engineering systems.

References

- Abbott A. 2001. *Chaos of Disciplines*. Chicago: University of Chicago Press.
- Abbott A. 1988. *System of Professions*. Chicago: University of Chicago Press.
- Biagoli M. 1999. *Science Studies Reader*. London: Routledge.
- Bartolomei 2007. *Qualitative Knowledge Construction for Engineering Systems: Extending the Design Structure Matrix in Scope and Procedure*. Engineering Systems Division. Cambridge MA, MIT PhD.
- Becker, H. S. (2001). *The Epistemology of Qualitative Research*. Contemporary Research Methods. R. M. Emerson. Long Grove, Illinois, Waveland Press, Inc Serman 1994
- Charmaz, K. (2001). *Grounded Theory*. Contemporary Field Research. R. M. Emerson. Long Grove, Illinois, Waveland Press.
- Charmaz, K. (2004). *Grounded Theory*. The Sage Encyclopedia of Social Science Research Methods. M. S. Lewis-Beck, A. Bryman and T. F. Liao. Thousand Oaks CA, SAGE. 2: 440-442.
- Dong, Q. (2002). *Predicting and managing system interactions at early phase of the product development process*. Mechanical Engineering. Cambridge MA, MIT. PhD.
- Glaser, B. and A. Strauss (1967). *The discovery of grounded theory*. Chicago, Aldine.
- Rein M. and Winship C. 1999. "The Dangers of 'Strong' Causal Reasoning in Social Policy," *Society*, Volume 36, No. 5 (July/August 1999), pp. 38-46.
- Sandelowski, M. (2004). *Qualitative Research*. The SAGE Encyclopedia of Social Science Research Methods. M. S. Lewis-Beck. Thousand Oaks, SAGE Publications. 3
- Serman J. D. (1989), "Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment", *Management Science*, 35(3), 321-339
- Serman, J. D. (1992) *Teaching Takes Off: Flight Simulators for Management Education*. *OR/MS Today*, 40-44.