# RELATIONSHIP BETWEEN THE LEVEL OF DETAILS AND UNCERTAINTY IN DESIGN RESEARCH

Meina Ke

A Thesis
In
The Department
Of
Concordia Institute of Information Systems Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master Applied Science (Quality Systems Engineering)

at Concordia University

Montreal, Quebec, Canada

November 2012

**CONCORDIA UNIVERSITY**
**School of Graduate Studies**

This is to certify that the thesis prepared

By: Meina Ke

Entitled: Relationship between the Level of Details and Uncertainty in Design Research

and submitted in partial fulfillment of the requirements for the degree of

**Master of Applied Science (Quality Systems Engineering)**

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

________________________________________ Chair

Dr. S. Li

________________________________________ CIISE Examiner

Dr. A. Awasthi

________________________________________ External Examiner

Dr. A. Akgunduz

________________________________________ Supervisor

Dr. Yong Zeng

Approved by ________________________________________________________

Dr. M. Debbabi, Director

Concordia Institute for Information Systems Engineering

_____________2012 ________________________________________________

Dr. R. Drew, Dean

Faculty of Engineering and Computer Science

## ABSTRACT

**Relationship between the level of details and uncertainty in design research**

Meina Ke

There are two main purposes of this thesis. First, it refines existed EBD coding scheme which used to be called formal model of EBD process and develops a three-layer structure of EBD coding scheme. The refined EBD coding scheme is expressed as hierarchical structure. It is composed by three levels. The first level includes collecting, processing and expressing. The second level describes the design process as five basic parts: identifying problems, searching information, generating solutions, expressing solutions and evaluating solutions. The third level is formed by 7 parameters: analyzing problems, identifying conflicts or (new) requirements, searching synthesis knowledge, searching evaluation knowledge, generating solutions, expressing solutions and evaluating solutions.

Second, it focuses on applying this coding scheme in a case study to find out the relationship between the level of details and uncertainties in design research. Through hypothesis test, the conclusion shows that the uncertainty increases when the level of details is deeper under the fixed technique. In this thesis, a designer was assigned to solve a design task, and the whole design process was recorded as a verbal protocol. The subjects apply the EBD coding scheme to define the actions of the design.

This coding scheme has several benefits. First, this coding scheme is a generic and logical. In the study, five subjects were chosen to do the experiments. Via EBD coding

scheme, the participants coded all the segmentations of the design process step by step. Second, the refined scheme is apt to understand and convenient for manipulating. The definition of the parameters of the EBD coding scheme is based on the previous researches and sources. No matter the subjects from which kind of background, through short-term trainings and continuous practice, they can operate the scheme easily and smoothly. The result of the protocol data gets a high percentage of agreements from different coders. Third, EBD coding scheme is a dynamic and developing system. The author can add more factors for building more levels. Fourth, encoding design process can further understand the cognitive thinking of a designer. The next object is expected to quantify and improve the design process and to get more perfect and complete design results.

In the study, it not only refines EBD coding scheme, but also applies the scheme in the same design protocol. According to the data analysis, we can get some results. Although the lower level can get a high agreement and low variance, this level cannot grasp adequate information from the design protocol. The higher could get a large amount of information. Nevertheless the percentage of agreement is low, and variance of the third level is high.

**Acknowledgments**

Numerous people attend and provide suggestions in the course of the study. Using the space here, I would like to express my appreciation to all the people whose have ever supported me to complete this thesis.

First of all, I would like to express my sincere gratitude to my supervisor, Dr Y.Zeng, for giving me an opportunity to complete this challenging and interesting research. His invaluable support and patient guidance constantly encourage me throughout the process of my research.

Then, I am grateful for Thanh An Nguyen for her assistance of designing experiments, her creative thinking influences me a lot throughout the research.

Also, I am glad to thank the other colleagues who participated to do the experiments. Thanks for taking time and paying attention to complete the experiment. Without the experiment data, I could not get result and finish the thesis.

At last, I would like to express my appreciation to my family. For their assistance, support, and encouragement, I finally could accomplish the research.

## Table of Contents

## List of Figures

## List of Table

# 1. Introduction

## 1.1 Background

Design is indicated by an agent, in a specific situation, by the means of basic tools to achieve goals which have to eliminate constraints and satisfy requirements. In general, the design is expressed by drawing or sketch. Recently, researchers have paid more attention on design processes rather than the notional design. They intend to fully understand what the designer thought and how the designer got the solution during the design process. Hence, analyzing the design process is crucial to achieve this goal. In the past, design research was considered as primarily research of design process. At present, the notion has been extended to pursue to understanding and improving the design process from different aspects, such as creative of design, designer, product, or behavior of designer. Based on the previous observations, when a designer tried to recall the design process, he or she could not remember everything in the design, even missed some significant information. Confronting this challenge, some tools should be used to keep a record for the whole design process. Protocol analysis plays an important role which helps record a design process in cognitive design. The procedures include recording design process, dividing the whole process into small units, encoding the design process and analyzing the process. During several decades, a number of coding schemes have been proposed and applied for analyzing the same design protocol. Most of the coding models are based on the observations of design. Using these coding schemes, researchers can collect large amounts of information of variety aspects. For instance, French represented a model that composed by four major activities: problem analysis, design conceptual, coding scheme embodied and detail (French, 1998). Cross proposed a model of design process. It included four factors that called exploring, generating, evaluating and communicating (Cross, 2008). Gero devised FBS (J. S. Gero & Kannengiesser, 2006) and situated FBS ontology (J. S. Gero & Kannengiesser, 2004). Gero used the coding schemes to capture semantic information from design protocol studies. Gero also proposed another scheme named “action categories” (Suwa, Purcell, & Gero, 1998). The action categories scheme divides the design activities into big parts: physical, perceptual, functional and conceptual. The relations between different parts are also coded. Shah (Hernandez, Shah, & Smith, 2010) identified key components and develops effectiveness metrics to understanding the cognitive mechanisms in design ideation. Dong (Dong, Kleinsmann, & Valkenburg, 2009) proposes computing the language appraisal in design. A long-term design project used by the University of Maryland chooses four variables to describe the design process: design step, information processed, activity, and object (Mullins, Atman, & Shuman, 1999).

However, using these coding schemes can help researchers extract a mass of information from the same design protocol, but the uniqueness of the each scheme make it hard to compare the same design protocol. As a consequence, a generic coding scheme is required to yield high consistence and unit results which is used to describe the whole design process and get further understanding of the cognitive thinking.

## 1.2 Objective

Although many of coding methods can be used at the same design protocol, it is still hard to compare the results of the protocol data. Hence, a refined EBD coding scheme is used to solve this issue (Nguyen & Zeng, 2012).

The objective of the thesis is to refine EBD coding scheme and apply the scheme for determining the level of details in design research. The refined coding scheme is generic, reliable and logical. It confirms that almost every activity related to the design processes can be defined. Through data analysis and hypothesis test (Spiring, 2007), the scheme is proved to get a high agreement of design actions.

The aim is to define cognitive action in the design processes and extract semantic information from the experiment protocol. According to the EBD coding scheme, the researchers can represent the behaviors of the design which helps further understand the designers' thinking processes, and better know how designers get the design solution.

The design protocols are conducted by colleagues in design lab of the Concordia University. The colleagues collect 22 design reports in all. In the study, the standard of choosing the protocol data is described as following:

(1). The time of the design process should be appropriate, not too long or too short.

(2). The design solution should be integrated which satisfying the requirements of the design task.

## 1.3 Contribution

In this study, the purpose is not only to use coding scheme to define activities, but also validating the relations between level of details and uncertainty. This scheme includes three levels. The first level is composed of three parameters: collect, process, express. There are five parameters in the second level: identify problem, search knowledge, generate solution, express solution and evaluate solution. Most of the verbs of action are concluded in the second level. Hence, in the third level, the author collects some nouns for extending the language system. Consequently, the third level includes seven elements: analyze problems, identify conflicts/requirements, search synthesis knowledge, search evaluation knowledge, generate solution, express solution, and evaluate solution.

Followed EBD coding scheme, coders define the actions which help describe the design process, and refine the definitions of action level by level. Three levels of EBD coding scheme are respectively independent. Based on hypothesis test, the author can gain the results as follows:

(1). The lower level gets a higher percentage of agreement and smaller variance, however, level 1 cannot obtain adequate information from the design process.

(2). On the contrary, the higher level gains a large amount of information. Nevertheless, it brings about a lower percentage of agreement and bigger variance.

The results do not mean that the lower level is better. In order to further understand the thinking process, the higher level is required. Hence, some techniques or tools should be used to decrease the variance and increase the percentage of agreement in the coding process.

### 1.4 Thesis organization

In the second part, the author gives an outline of elements in protocol analysis. Several coding schemes are represented as well. In the third and fourth parts, the author depicts the structure of the EBD coding scheme, the application of the coding scheme is shown in a case study. In the fifth part, hypothesis test is applied to analyze the protocol data and get a series of results. The last part introduces a brief interpretation of the currently situation and future development.

The organization of the thesis is shown in **Error! Reference source not found.**.

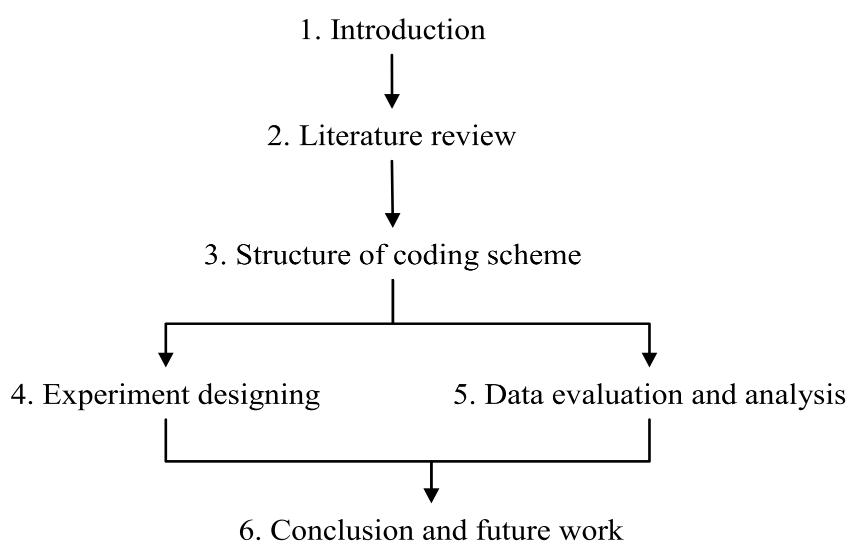


**Figure 1.1 Organization of the thesis**

# 2. Literature Review

## 2.1 Introduction

Design can be defined as a process of creating artifacts or solving problem to satisfy the requirements of the tasks (Jin & Chusilp, 2006). In order to reach the objects of the tasks, the designers have to figure out requirements and solve conflicts. In several decades, researchers focused on the "thinking process" of the design. Protocol analysis plays an important role in understanding and analyzing the designers' thinking process. Generally, the protocol analysis includes several parts: verbalization, segmentation, encoding, and data analysis (Anders & Simon, 1980). This chapter concludes a brief review of research in variety coding mechanisms and segmentation methods in a design protocol.

## 2.2 Protocol analysis

At the very beginning, protocol analysis is considered as a psychological method which used to study thinking in cognitive science and cognitive psychology. At present, this tool is broadly applied in different domains. Protocol analysis is a widely used research method which collects verbal reports from participants in design (Hughes & Parkes, 2003). Using techniques of protocol analysis, researchers can comprehend the designers' cognitive process. Some techniques such as interviews and surveys are usually used to gain information in the design process which cannot observe directly. Nevertheless, these techniques are criticized by recoding uncompleted report. Even worse, the interviews lead to inaccurate reports.

Alternative technique, verbal protocol analysis is widely used which is suitable to diverse types of problems. For instance, verbal protocol analysis is applied to different domains which include cognitive science (Simon & Kaplan, 1989)，artificial intelligence (Conati & Vanlehn, 2000), human-computer interaction (Howard, 1997) and behavior & information technology (Herbsleb et al., 1995) and so on. According to previous studies, the verbal protocol analysis is generally divided into four basic parts: transcription, segmenting, encoding, and statistical analysis. The phrases of the protocol analysis can be represented in

Figure 2.1. For the next step, the study introduces every part respectively.

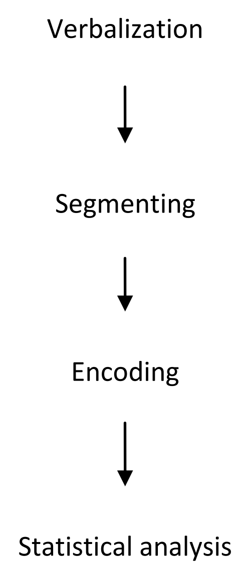


**Figure 2.1 Phases of protocol analysis**

### 2.2.1 Verbalization

First step of protocol analysis is recording the design process. Three types of verbalizations are indicated as follows:

(1). The vocalization is simple which is possible without further processes, the information is reappeared directly (Pennington, Nicolich, & Rahm, 1995).

(2). A verbal form is used to record the information (Moore & Lehman, 1995).

(3). The information requires further processes (Lemut, Dettori, & Boulay, 1993)(Chi, Bassok, Lewis, Reimann, & Glaser, 1989).

Think aloud method belonged to first type which will not change the structure of designers' thought processes. It is extensively used in generating verbal protocols. Subjects found it is difficult to remind complete thought process. Therefore, the verbal reports are recorded with video-tape. The actions of the designer contribute to more information, especially in some environments.

### 2.2.2 Segmenting

After collecting the verbalization data, the design process will be divided into small unites which are named as segmentation. Segmentation is considered to constitute an independent process. There are two kinds of segmentation methods.

One principle of segment is based on the pause, intonation and syntactical markers (Hayes, 1986). In this study, the protocol data is recorded from video protocol, the author divided the design protocol based on pause and syntactical markers. After segmentations are identified, they are encoded by random coders. Hence, it is important that and the context of the segment should be as small as possible. In addition, the segmentation is large enough to permit independent coding.

The other way of segmenting verbal reports is based on designers' intention. A new segmentation starts when the intention of the subjects or content of thoughts changes. The segmentation is possibly composed by one sentence, sometimes by many. However, the method of segmentation brings about ambiguity when distinguishing the diverse intentions. This method increases obstacles when the coders translate the design protocols.

## 2.3 Coding mechanisms in protocol analysis

Encoding is not only the third phase of protocol analysis, but also the main objective of this study. In this part, several coding mechanisms are introduced which are widely used and work well in different domains.

### 2.3.1 FBS and situated FBS ontology

In FBS framework, three variables are used to describe a design object with different aspects: function, behavior, and structure.  There are eight processes which link the function, structure and behavior together. In Figure 2.2, it shows that the direct connection between function and structure is not existed. The set of processes include formulation, synthesis, analysis, evaluation and documentation.

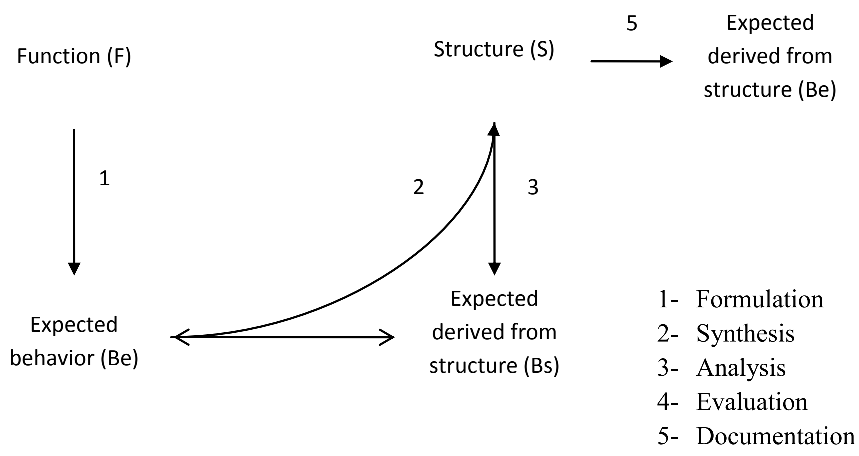


**Figure 2.2 The FBS ontology**

There are other three reformulations which are not shown in the figures. In the FBS ontology, these reformulations were considered as the most remarkable processes. The definitions of the three reformulations are shown in Table 2.1.

**Table 2.1 Definitions of three reformulations**

| Number | Description |
|---|---|
| Reformulation type 1 (process 6) | If the real activities do not satisfy the objects, some changes related to structure variables are addressed. |
| Reformulation type 2 (process 7) | If the real activities do not satisfy the objects, some changes related to behavior variables are addressed. |
| Reformulation type 3 (process 8) | If the real activities do not satisfy the objects, some changes related to function are addressed. |

In the FBS framework, environment is considered as static. In fact, through the whole design process, the context is dynamic and changing. Therefore, the situated FBS framework is proposed which based on extended the previous framework. In the situated FBS framework, the context is divided into three parts: expected world, interpreted world, and external world. Due to the dynamic environment, the eight basic processes are extended to twenty.

### 2.3.2 Action categories

In 1998, Masaki devised a new scheme for coding designers' activities. Designers' action was divided into four categories: physical, perceptual, functional and conceptual. In Figure 2.3**Error! Reference source not found.**, the parameters in different categories are expressed.

Using this coding scheme, we can roughly describe the design process. The limitation of this scheme is that some segments are too ambiguous of code G-actions with a unique description.

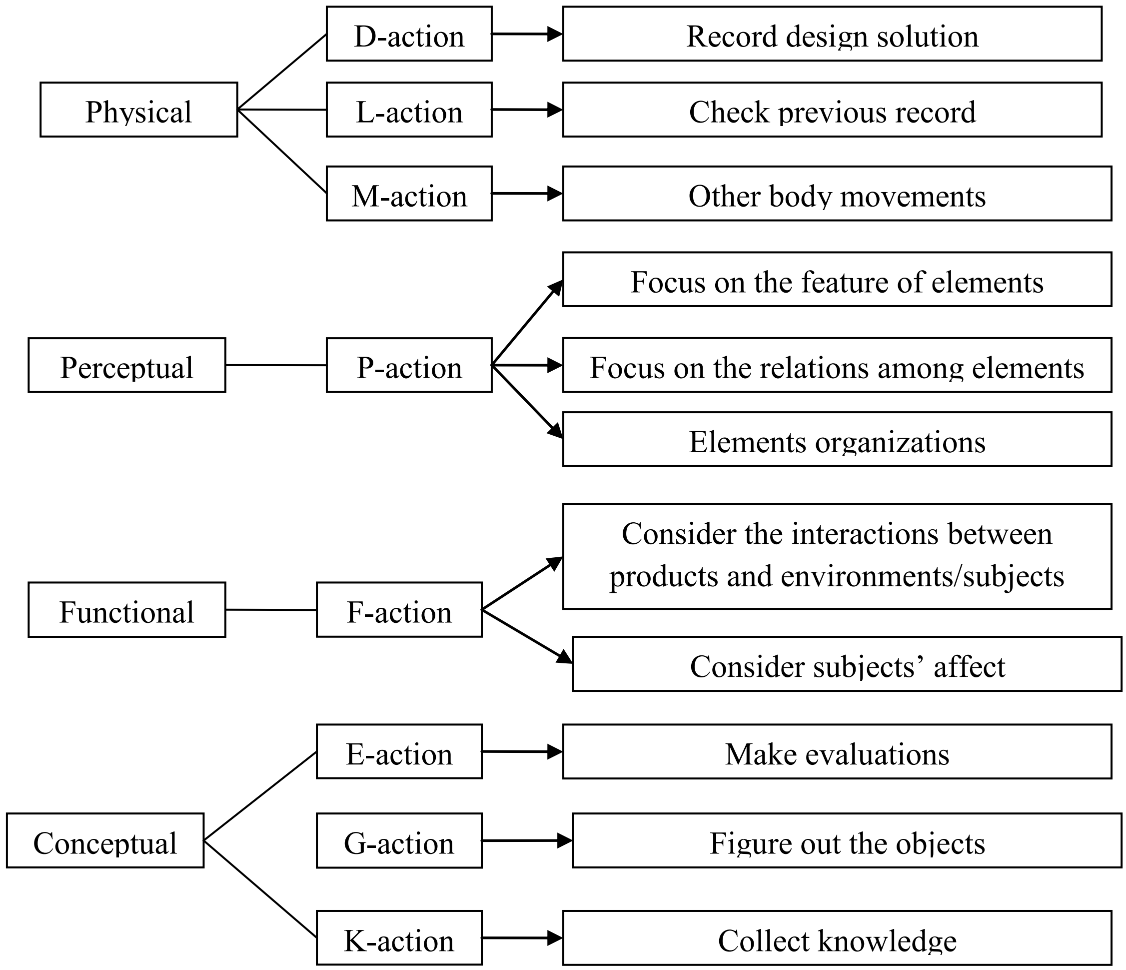


**Figure 2.3 Action categories**

### 2.3.3 Language of appraisal

Growing design practices suggest that how a designer's "feeling" influences the way that a designer behaves. In 2006, Dong proposed a linguistic system of appraisal (Andy Dong, 2006)(Wang & DONG, 2007). Andy Dong also focused that the valence of the affective influenced the orientation of the linguistic appraisals (Dong et al., 2009). Appraisal is a representation through language of favorable and unfavorable attitudes towards specific subjects. The categories of appraisal include process, product and people. The structure of the categories is shown in Figure 2.5.

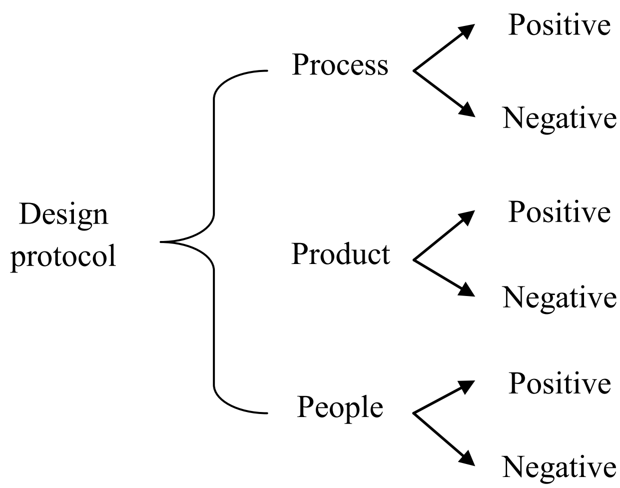


**Figure 2.5 Categories of appraisal**

The language of appraisal is building up by appraisal of three categories. In the appraisal of process, if the appraisal is based on the experience or personal explanation, then it is considered as appreciation. Conversely, if the appraisal is identified by norms, it is considered as judgment. Appraisals of product are one way which designers give subjective estimations. In the appraisal of people, the designer assesses others or himself subjectively. In order to restrict the potential scope, the appraisal of people was divided into four parts: affect, cognitive, cognitive-behavioral and capability. Hence, the framework of the linguistic appraisal was shown in Figure 2.6.

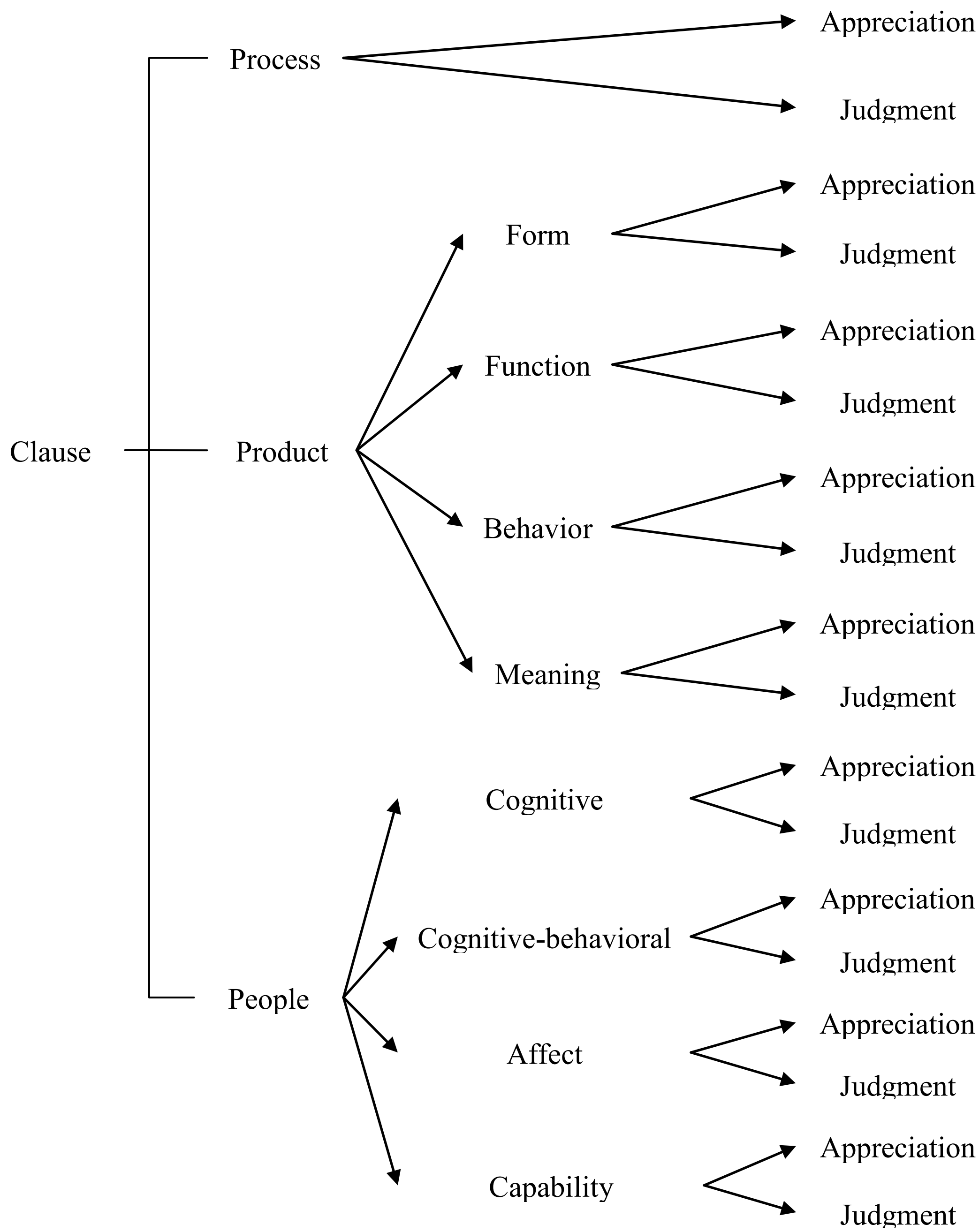


**Figure 2.6 Structure of the language of appraisal in design**

### 2.3.4 The Structure-of-Intellect (SI) model

In 1971, Guilford set up the structure –of-intellect (SI) model (P, 1956). In a long period, SI Model was widely used in multiple domains. The SI model is a morphological model. It is composed of three parts: operation, content, and product. Further, these parts are divided into different parameters. The structure of SI model is indicated Table 2.2.

**Table 2.2 Structure of SI Model**

| Operations | Content | Products |
|---|---|---|
| Cognition | Figural | Unit |
| Memory | Symbolic | Class |
| Divergent production | Semantic | Relation |
| Convergent production | Behavioral | System |
| | | Transformation |
| | | Implication |

According the model, there were 120 kinds of action in all. In 1988, Guilford updated the structure of intellect (SI) model (Guilford, 1988). First, figural–content factor was divided into two categories: visual and auditory. Second, the memory-content factor was separated into memory retention and memory recording. Consequently, there are 180 kinds of activity in total. The structure of the revised SI model is expressed in Table 2.3.

**Table 2.3 factors of the revised SI model**

| Content | Products | Operations |
|---|---|---|
| Visual | Units | Evaluation |
| Auditory | Classes | Convergent production |
| Symbolic | Relations | Divergent production |
| Behavioral | Systems | Memory retention |
| Semantic | Transformations | Memory recording |
| | Implications | Cognition |

The limitation of the revised SI model was obviously existed. A huge challenge was to distinguish the diversities of activities which will increase difficulty to use this coding scheme to identify designers' actions.

### 2.3.5 Generic model

In general, protocol data is analyzed by individuals. In fact, it could be analyzed by groups as well. In 2002, a generic step model of design team activities was proposed by Petra (Stempfle & Badke-Schaub, 2002).

This Generic Model is comprised of three factors: content, cognitive operation, process. Petra agreed with Ward who established Generic model that generation and exploration are essential ingredients in a design process, but only two operators are not enough to solve problems. These two operators can help to broad possible solutions. So, they proposed comparison and selection operators to narrow down the problem space. Consequently, the basic thinking operation includes four elements: exploration, generation, comparison and selection. Regard to content which belonged to goal space, it was divided into six phases: goal clarification, solution generation, analysis, evaluation, decision and control.

Process action is related to solution space. It includes five steps: planning, analysis, evaluation, decision, control.

This model enables to decompose the complex design activity into small chunks which can by analysis by a variety of methods or tools. Thus, it provides a precise picture of what the designers really do.

The structure of the generic model is shown in Figure 2.7.

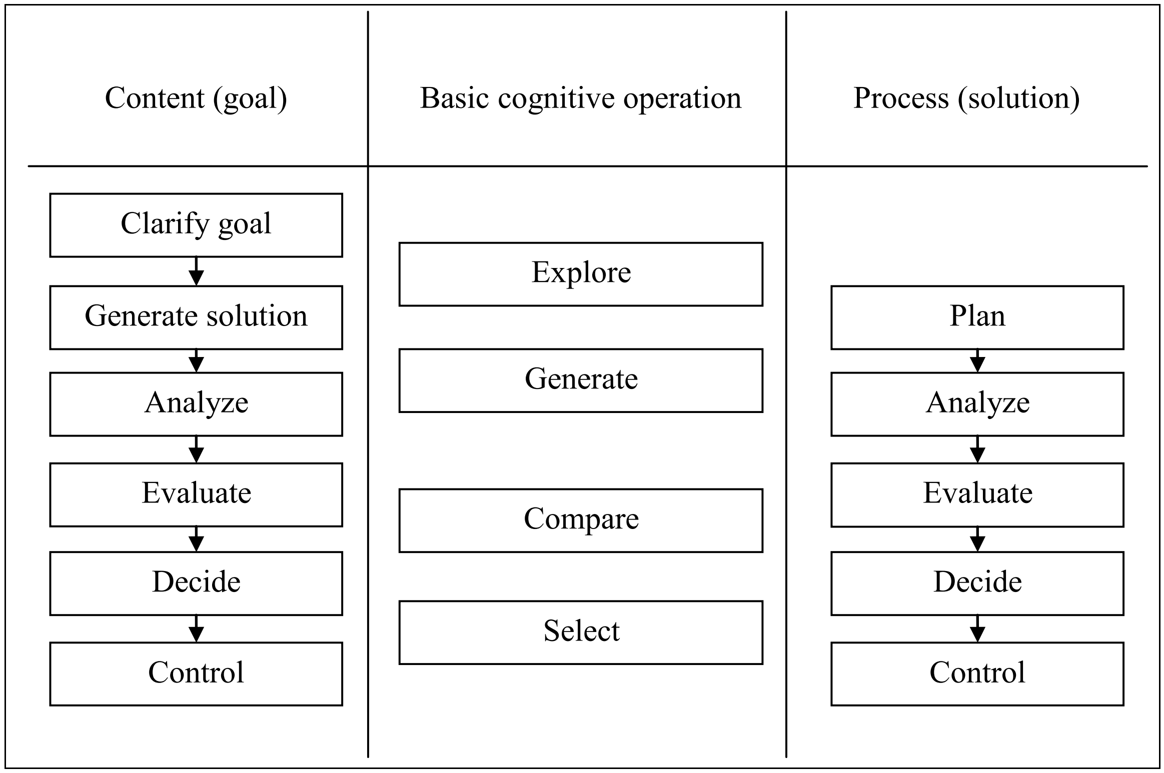


**Figure 2.7 Generic step model of design team activities**

However, this model is usually used in design teams rather than by an individual designer. It focuses on communications among designers in a group. When applying this model, designers have to modify solutions to develop a satisfying or optimal solution through iterative processes.

### 2.3.6 Generic design activities

Sim insisted that there was not a consensus definition of the activities during the design process. Therefore, an identification and classification of generic design activities (Sim & Duffy, 2003) was proposed to reach an agreement of understanding of these activities. This ontology was based on previous published literature. It classified the activities by design definition, evaluation and management. The remarkable contribution of this model was building an effect design support system and reusing design. All the parameters are listed in Table 2.4.

**Table 2.4 Identification and classification of generic design activities**

| Design definition activities | Design evaluation activities | Design management activities |
|---|---|---|
| Abstracting | Analyzing | Constraining |
| Associating | Decision making | Exploring |
| Composing | Evaluating | Identifying |
| Decomposing | Modeling | Information gathering |
| Defining | Selecting | Planning |
| Detailing | Simulating | Prioritizing |
| Generating | Testing / experimenting | Resolving |
| Standardizing | | Searching |
| Structuring / integrating | | Selecting |
| Synthesizing | | Scheduling |

### 2.3.7 Cognitive activity model

Jin proposed a cognitive activity model (Jin & Chusilp, 2006) which was used to expound the thinking steps of design process. In this model, it did not focus on capturing or processing information. The focus was on depicting steps or works of the whole design process. The key activities of the cognitive model are represented in Figure 2.8. In the cognitive model, three global iteration loops existed which were not shown in Figure 2.8. The iteration loops are divided into two classifications: information flows and activities. These loops inserted among the four major activities, they were defined as problem redefinition loop, idea stimulation loop and concept reuse loop.

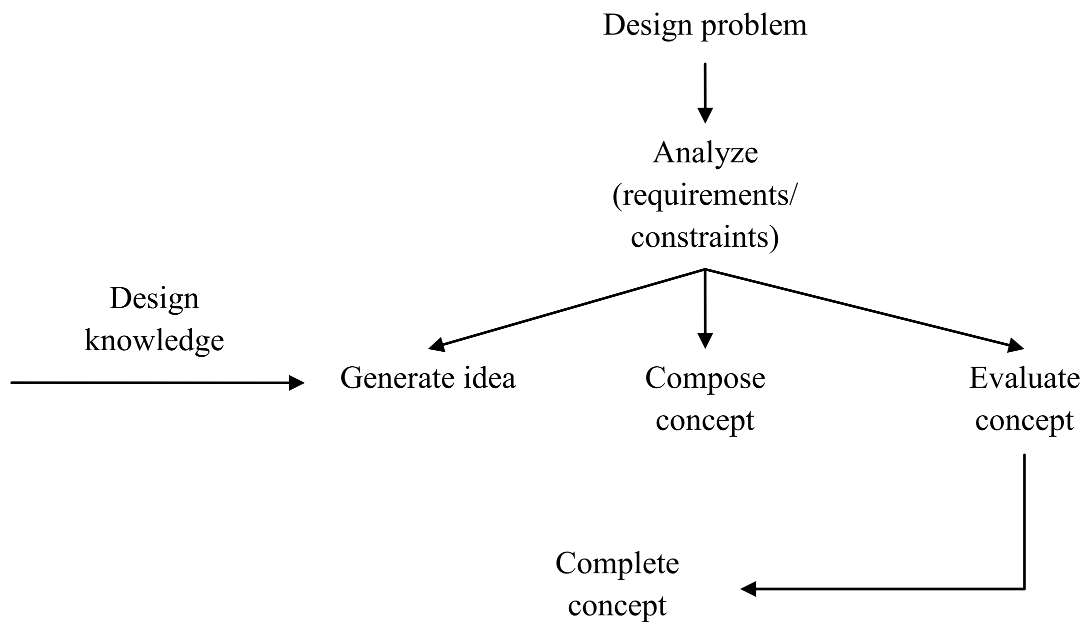


**Figure 2.8 Cognitive activity model of conceptual design**

## 2.4 Current limitations

There are a variety of coding mechanisms that can be used in the design protocol. These coding schemes can help to get a rich understanding of information in different aspects in the same design protocol and each scheme has its own uniqueness and diversity.

We encounter a problem that it is difficult to compare the results in the same protocol. Therefore, we need to devise a generic coding scheme that can generate consistence results. The uniform results describe the behavioral of the designers, and contribute to provide a deeper understanding of the design thinking and activities.

# 3. Structure of EBD Coding Scheme

## 3.1 Introduction

Previous chapters mention that coding the segmentations of the protocol data is a critical part of the study. Consequently, building a coding system to analyze the design protocol is significant. In this study, the refined EBD coding scheme is generic and logical. The results of the data are reliable. Meanwhile, using this coding scheme, the coders gain a uniform result in the same protocol. This is also the main purpose of this thesis.

## 3.2 Environment Based Design (EBD)

Design is considered as a process which stems from environment, works for environment, and finally changes the existed environment to reach a purpose (Zeng, 2011).

Environment based design (EBD) is a methodology which derived from observation of designs. It is a logical and recursive process that aims to provide designers the right direction for solving a design problem. It includes three activities: environment analysis, conflict identification, and solution generation. The three activities work together to update environment and its internal relationships to solve a design problem. The design process continues with new environment analysis until no more undesired conflicts exist.

In order to conduct the design direction, a natural graphic language called Recursive Object Model (ROM) was proposed to analyze the existed circumstances (Zeng, 2008). By the means of ROM, designers can divide a design problem into small objects, through analyzing the relationship between different objects. All the conflicts and requirements can be represented. When the last requirement is solved, the design is supposed to finish.

Generally, a design problem and a design solution are expressed by hand-writing or sketches in a design process.

Via ROM and following the steps of the EBD method, the goals or requirements of the design can be totally completed. A detailed formal model is used to describe the design process which can be found in **Error! Reference source not found.**. The detailed formal model depicts a basic pattern how the designer uses information to solve the design problem.

### 3.3 Coding scheme

As mentioned in the previous part, EBD is a method which guides the designer follow a right direction to solve the problem step by step. The formal model (Nguyen & Zeng, 2012) describes the detailed process of design. The model of design process is shown in Table 3.1. In order to simplify the formal model, the author builds up hierarchical structure to refine the model. This hierarchical structure helps analyze the design process level by level, and refine the actions from roughly to minutely. In this thesis, three levels of scheme are presented.

**Table 3.1 EBD formal process**

| Activity | Algorithm | Description |
|---|---|---|
| 1 | $r_i^{ad} \in R_{i-1}^d \cap R_a^d$; | Identify a critical requirement $r_i^{ad}$ from a list of requirements $R_i^d$ to start the design |
| 2 | $\forall r_i^{ad}, \exists K_a^s (K_a^s : r_i^{ad} \rightarrow S_a^i)$; | Search for the right synthesis knowledge $K_a^s$<br>Generate tentative primitive design solution $S_a^i$ |
| 3 | $\forall S_a^i, \exists (r_i^{ad})', (r_i^{ad})^+ = (r_i^{ad}) \cup (r_i^{ad})'$; | Search new design requirements $(r_i^{ad})'$ based on the tentative primitive design solution $S_a^i$<br>Update design requirements to $(r_i^{ad})^+$ |
| 4 | $\forall S_a^i[P_a], \exists K_a^p, ((K_a^p = [S_a^i \cup [P_a], L]_{\approx}) \cap (K_a^p : S_a^i \rightarrow P_a))$; | Search for the knowledge $K_a^p$ based on design solution $S_a^i$ and design requirements $[P_a]$<br>Derive the performance $P_a$ of the primitive design solution $S_a^i$ |
| 5 | $\exists S^{[s]} \in S_a^i, \exists P^{[p]} \in P_a, K_a^e(S^{[s]}) \cap K_a^e(P^{[p]})\ if\ S^{[s]} = \emptyset \vee P^{[p]} = \emptyset, go\ to\ 2$; | Validate if the primitive solution $S_a^i$ meets structural and performance requirements |
| 6 | $\forall P^{[p]}, \exists S^{[p]} \in S_a^i, (K_a^p : S^{[p]} \rightarrow P^{[p]})\ if\ S^{[p]} = \emptyset, go\ to\ 2$ | Verify if knowledge exists to evaluate the performance of the primitive solution $S_a^i$ |
| 7 | $S' = S_a^i \cap S^{[s]} \cap S^{[p]}\ if\ S' = \emptyset, go\ to\ 2$; | Identify a common solution $S$ based on those from performance and structural requirements. |
| 8 | $S_p^i = \xi(S', S_p^{i-1})$; | Add the newly generated primitive solution $S'$ to already completed intermediate solution $S_p^{i-1}$. |
| 9 | $X_a^i = K_a^p(S_a^i); X_p^i = K_a^p(S_p^i)$; | Identify the right performance knowledge $K_a^p$<br>Analyze the performance of the newly generated primitive design solution $S_a^i$ and the existing partial design solution $S_p^i$. |

| | | |
|---|---|---|
| 10 | $R' = X_a^i \updownarrow X_p^i = K^e\left(\left(X_a^i \otimes X_p^i\right) \cup \left(X_p^i \otimes X_a^i\right)\right);$ | Search for the conflicts $R'$ between the performances $X_a^i$ *and* $X_p^i$ of the newly generated design solution $S_a^i$ and the previously generated partial design solution $S_p^{i-1}$ |
| 11 | $R_i^d = (R_i^d/\{r_i^{ad}\}) \cup R';$ | Redefine the design requirements |
| 12 | $if\ R_i^d \neq \emptyset, go\ to\ 1\ ;$ | Stopping condition |
| 13 | $S = S_p^i;$ | Output the design solution. |

### 3.3.1 The first level of coding scheme

#### 3.3.1.1 Process description

Design aims to change the current environment to a desired one by generating a new product. Three basic activities happen during the design process: environment analysis, conflict identification, and solution generation.

In a design experiment, three major factors work together: designer, product and environment. In order to generate a solution or product for a design task, the designer should collect a large amount of information from environment, and the change of the environment will affect the solution frequently. In a general design process, people who solve the problem first collect the information from the environment. According to process the information, the previous environment is changed to satisfy the new product until the all the problems have been solved. The whole process is concisely explained in **Error! Reference source not found.**.

#### 3.3.1.2 Parameter definition

The three key words which are chose from the design process compose the first level. Three parameters are respectively expressed as collect, process, express.

- Collect: gather and extract information from a number of sources.
- Process: integrate collected information so that solution is generated.
- Express: after generating the solution in the mind, draw or write down the solution by hand.

The process of level can be shown as Figure 3.1.


**Figure 3.1 Flow chart of the first level**

#### 3.3.1.3 Parameter definition

The three key words which are chosen from the design process compose the first level. Three parameters are respectively expressed as collect, process, express.

- Collect: gather and extract information from a number of sources.
- Process: integrate collected information so that solution is generated.
- Express: after generating the solution in the mind, draw or write down the solution by hand.

### 3.3.2 The second level of coding scheme

#### 3.3.2.1 Process description

In the formal model, when the designer firstly gets the design task, he or she should figure out the objects or the problems of the design. Confirming the problem of the design, information related to the problem has been collected. Gaining adequate of the information, the solution is generated and expressed. Before finishing the whole design, the solution has to been evaluated. If the design solution solves the design problems, the design process is completed; if not, the designer should go back to first step of the design. The whole design process can be described as Figure 3.2.

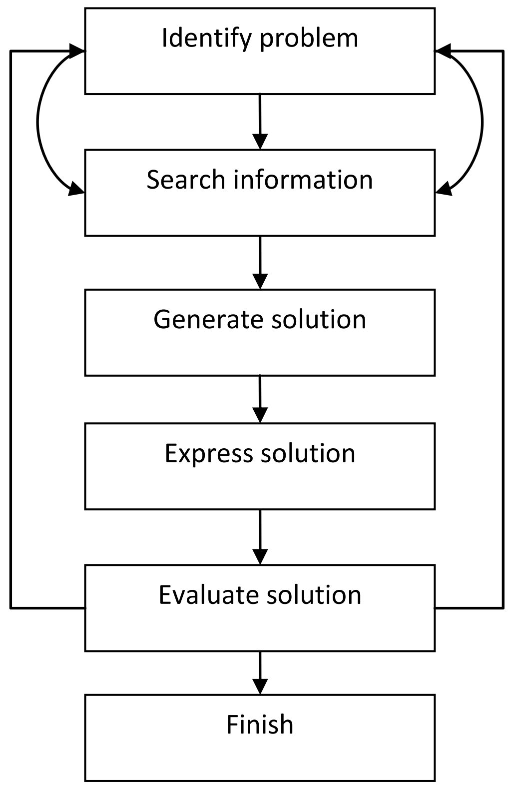


**Figure 3.2 Flow chart of the second level**

#### 3.3.2.2 Parameter definition

Several main words can approximately describe the whole design process. The five parameters of the second level include identifying, searching, generation, expression, evaluation.

- The definitions of five parameters are indicated as follows:

- Identify: based on the problem, understand and pick up information from a number of resources.
- Search: collect information from multiple resources, such as network, experiences, background.
- Generate: after searching information for the design task, generating is what you think in your mind. Generating is a process from nothing to something for the solution.
- Evaluate: asses if the generated design concept is useful, good or reliable.

### 3.3.3 The third level of coding scheme

#### 3.3.3.1 Process description

In design, a good solution is measured by if the solution can satisfy all requirements. In EBD, a tool called ROM is used to list all the requirements. The designer firstly analyzes design problems, starts the design with identifying a list of requirements and conflicts. For the next step, he or she searches synthesis knowledge to generate solution, expresses the solution by writing. Based on the primitive solution, designer searches evaluation knowledge to evaluate solution. If the solution is good, the designer finishes the design; if not, the designer should go back to analyze problem or search synthesis knowledge for new requirements (Zeng & P.Gu, 1999). In third level, the types of the knowledge are refined in order to gain more details and information of the thinking process. The process of the third level is shown in Figure 3.3.

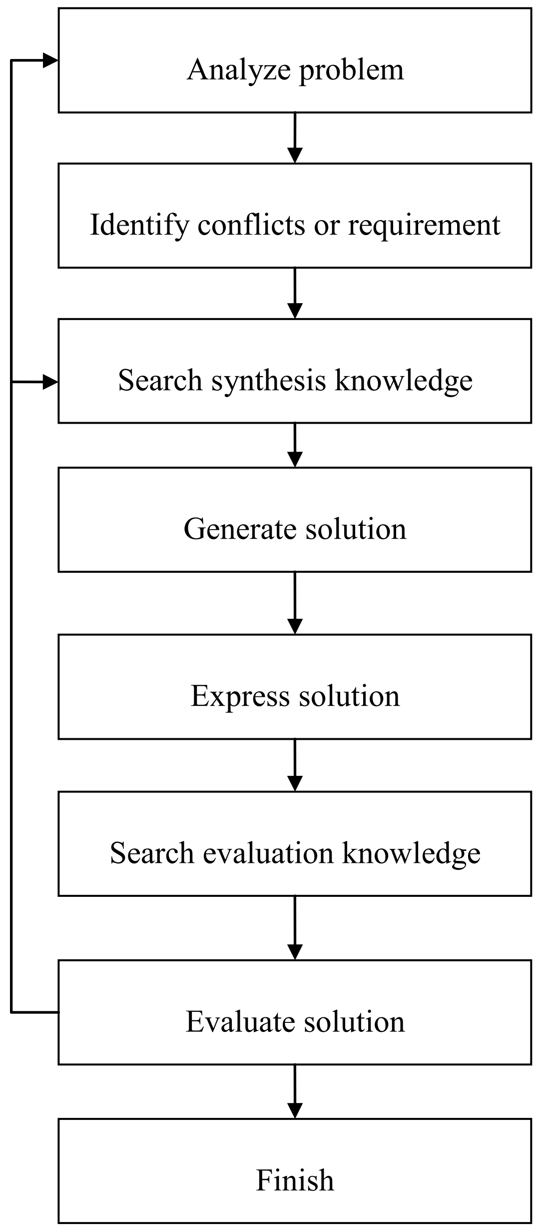


**Figure 3.3 Flow chart of the third level**

#### 3.3.3.2 Parameter definition

There are four different kinds of knowledge: requirements, conflicts, synthesis knowledge, and evaluation knowledge.

- Requirements: purposes of the design, generally getting from the design tasks.

- Conflicts: problem between the requirements and solutions or between requirements.
- Synthesis knowledge: relation from requirement to solution
- Evaluation knowledge: relation from solution to requirement

The author also adds a verb in the third level to refine the behaviors in the design processes.

- Analyze: figure out the objectives or purposes of the design.

Consequently, the factors of EBD coding scheme are represented in detail in Table 3.2.

**Table 3.2 Structure of EBD coding scheme**

| level | Parameters |
|---|---|
| 1 | Collect, process, express |
| 2 | Identify, search, generate, express, evaluate |
| 3 | Analyze problem, identify conflicts/requirements, search synthesis knowledge, search evaluation knowledge, generate solution, express solution, evaluate solution |

## 3.4 Applied coding scheme into a design protocol

Understanding the meaning of parameters in each level is facile, but how to operate the coding system to define the cognitive action is the main problem. As everyone knows, the brain is a complicated human organ. It can dispose large amounts of information simultaneously, even in a short time. Therefore, it is very difficult to separate an independent action. In addition, in a design process, relations between actions are existed, such as dependencies and trigger relations.

Consequently, when coders use this scheme, the participants can choose more than one elements of the EBD coding scheme to define the segmentation. In order to simplify the experiment, the subjects are demanded to eliminate the action which is impossible to happen.

In this thesis, when the coder define the action by coding scheme, they need to consider the context, and choose the action which impossible to happen and then remove or cross the impossible actions. This small experiment is also considered as an inspection if the coders use this scheme to define the cognitive actions logically and smoothly. There is no a standard to measure the results. The only criterion is that the defined actions can completely describe the design process, and get the main information of cognitive thinking.

**Table 3.3 Segmentation of sample**

| # | description |
|---|---|
| 1 | The designer looks the design question for more than twenty seconds. |
| 2 | He draws three boxes and a mesh over the boxes. |
| 3 | He stops for about 5 seconds. |
| 4 | He chooses the red color (for the pen). |
| 5 | He writes something and erases it. |
| 6 | He stops for 9 seconds. |
| 7 | He chooses the black color. |
| 8 | He added doors for each boxes and draws two more boxes between the three boxes. |
| 9 | He stops for 3 seconds. |
| 10 | He chooses the red ink and then he stops for 3 seconds. |

In order to explain the operation distinctly, part of the verbal report of a sample protocol data is collected in Table 3.3.

The result of the sample protocol is indicated in Table 3.4 which shows the possible actions in segmentation. Standard answer is not existed, the subjects have to define the actions based on the comprehending of the coding scheme, experience, and knowledge.

**Table 3.4 results of coding part of sample protocol**

| # | The first level | The second level | The third level |
|---|---|---|---|
| 1 | Express | Express solution, evaluate solution | Search evaluation knowledge, express solution, evaluation solution |
| 2 | Collect, process | Identify problem, search information, generate solution, evaluate solution | Analyze problem, identify conflicts/requirements, search synthesis knowledge, search evaluation knowledge, generate solution, express solution, evaluate solution |
| 3 | Collect, express | Identify problem, generate solution, express solution, evaluate solution | analyze problem, identify conflicts/requirements, search evaluation knowledge, generate solution, express solution, evaluate solution |
| 4 | Collect | Identify problem, search information, evaluate solution | Analyze problem, identify conflicts/requirements, search synthesis knowledge, search evaluation knowledge, evaluate solution |
| 5 | Collect, process | Identify problem, search information, generate solution, evaluate solution | Analyze problem, identify conflicts/requirements, search synthesis knowledge, search evaluation knowledge, generate solution, evaluate solution |
| 6 | Collect, express | Identify problem, generate solution, express solution | Analyze problem, identify conflicts/requirements, search evaluation knowledge, generate solution, express solution, evaluate solution |
| 7 | Collect | Identify problem, search information, evaluate solution | Analyze problem, identify conflicts/requirements, search evaluation knowledge, evaluate solution |
| 8 | Collect, process | Identify problem, search information, generate solution, | Analyze problem, identify conflicts/requirements, search synthesis knowledge, search |

| | | | |
|---|---|---|---|
| | | evaluate solution | evaluation knowledge, generate solution, evaluate solution |
| 9 | Process | Identify problem, generate solution, express solution, evaluate solution | Analyze problem, identify conflicts/requirements, evaluation knowledge, generate solution, express solution, evaluate solution |
| 10 | Process, express | Identify problem, search information, evaluate solution | Analyze problem, identify conflicts/requirements, search evaluation knowledge, evaluate solution |

# 4. Design of Experiment

## 4.1 Introduction

In this thesis, subjects are assigned a design task which participants need to encode the same design protocol by EBD coding scheme. In this chapter, the preparations of the experiment are introduced such as choosing the subjects, setting up the procedure during the experiment, collecting the final data of the verbal protocol, and so on. Ultimately, all of the individual results were collected and organized for the next step of the study which analyzed the data for quantifying the agreements for the activities in the same design.

## 4.2 Selection of subjects

In this thesis, the only criterion for collecting the subjects is to understand English well. The coding scheme is created in English. Therefore, understanding the meaning of the parameters in the coding scheme is the fundamental of applying the scheme. In addition, the subjects should have a shorting training to inspect if the subjects have understood the coding parameters and grasped and applied the scheme to define the design activities of design process successfully.  In the experiment, five subjects are chosen randomly. In this experiment, these five participants are all from engineer department, however, they never know coding scheme before, and all of them need a short training before the real experiment to ensure the experiment results reliable.

## 4.3 Materials

The design task was chosen from 22 design protocols. The design question is as follow: "Design a house can fly from one place to another."

It was chosen because the process of the design is reasonable and smooth, the design solution is complete, and the length of the design is neither too long nor too short.

The verbal report is elicited from the video design protocol. In the previous part, two kinds of segmentation methods are introduced. In this study, segmentations are based on the pause and syntactical markers. This segment mechanism is in order to eliminate the fuzzy parts during the design process, and to separate the activity as independently as possible which will help coders to define the cognitive actions. This study is focus on applying the EBD coding scheme in a design search, hence, the author segmented the verbal report before experiment. Hence, the subjects just need to define the design activities based on the segmentations. The sketch which is shown in Figure 4.1 is the verbal report of the design experiment.

This experiment sketch includes the figure of the products and interpretation of the functions of different parts of the products. By the means of the sketch, the coders deeply comprehend the cognitive thinking of the design, further realize how to use EBD coding scheme to encode the activity of the design protocol.

In the experiment, the video that recoded the design process from different aspects of camera is also offered.

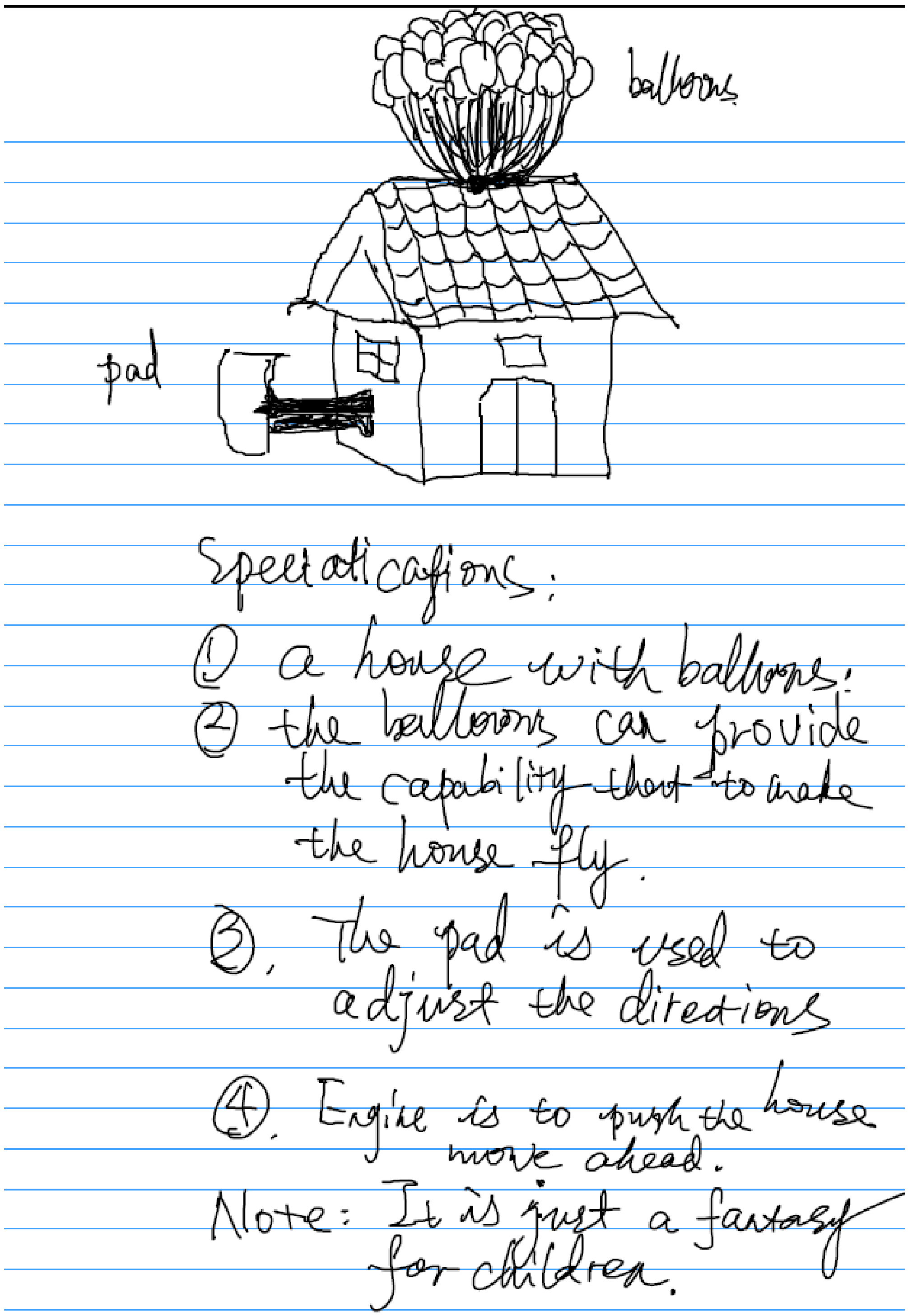


**Figure 4.1 Experimental sketch**

The part of segmentation results are listed in Table 4.1

**Table 4.1 Segmentations of design process**

| # | Duration (s) | Description |
|---|---|---|
| 1 | 7.42 | He looked the design question again and then minimized the page. |
| 2 | 12.82 | He stopped for a while. |
| 3 | 16.86 | He began to draw the house. |
| 4 | 6.12 | He stopped for a while. |
| 5 | 21.48 | He continued to draw the house. |
| 6 | 2.38 | He stopped for a while. |
| 7 | 6.25 | He erased the line on the roof. |
| 8 | 8.46 | He stopped for a while. |
| 9 | 6.31 | He used pen to connect the gap between the two short lines. |
| 10 | 3.05 | He stopped for a while. |
| 11 | 8.27 | He erased part of the roof. |
| 12 | 25.26 | He adjusts the device system. |
| 13 | 12.53 | He continued to draw the house. |
| 14 | 3.39 | He stopped for a while. |
| 15 | 6.47 | He erased some no use lines. |
| 16 | 6.31 | He seemed to find some tools. |
| 17 | 4.95 | He moved the page up and down. |
| 18 | 15.33 | He added a door for the house. |
| 19 | 0.92 | He stopped for a while. |
| 20 | 12.04 | He drew a window for the house. |
| 21 | 1.07 | He stopped for a second. |
| 22 | 15.87 | He drew some lines on the roof. |
| 23 | 2.74 | He stopped for a while. |
| 24 | 4.78 | He eliminated the last three lines on the roof. |
| 25 | 4.75 | He stopped for a while. |
| 26 | 7.47 | He added another three lines on the roof. |
| 27 | 1.68 | He stopped for a while. |
| 28 | 12.06 | He added some curves between two lines on the roof. |
| 29 | 2.45 | He connected the gap between two short lines. |
| 30 | 1.49 | He stopped for a while. |
| 31 | 5.13 | He added two lines on the side face of the roof. |
| 32 | 2.65 | He stopped for a while. |

## 4.4 Experimental method and data collection

### 4.4.1 Experiment procedure

During the experiment, several procedures are building up and have to obey throughout the whole experiment. The principles are indicated as follows:

(1). Find five persons to accomplish the experiment.

(2). The author looks for a quiet place to complete the experiment. Make sure the environment with the least noisy so that the participants will not be disturbed during the experiment.

(3). Set up the equipment; briefly explain the objective, procedures about the experiment to the participants.

(4). The author gives 30 minutes to the participant to understand the criteria of the experiment. If subjects do not understand or have some problems, they can seek assistance from the author.

(5). Take a min test for every participant to make sure they clearly figure out how to apply criteria to the experiments.

(6). As all the preparations are over, the participants could begin to conduct the test. There is no time limit. During the test, the participants can inquire the questions about the design at any time. In addition, the participants also can stop to have a short rest when they need.

(7). After all the experiments are completed, the author checks all the results.

### 4.4.2 Coding method

The design protocol is defined by the three levels. Hence, the verbal protocol needs to encode three times for each subject. The subjects can finish the whole design in one day or separate the experiment into several parts, and finish the experiment in a few days. It depends on the subjects' desire. Whereas, the subjects are demanded to at least accomplish the experiment by one level of scheme in one time.

### 4.4.3 Data collection

There are five subjects attended the experiment. Every subject encoded the design protocol three times. After all the experiments are finished, the author should collect all the experiment data and make a comparison between two different subjects. Finally, there are ten comparisons between five participants, three groups of comparison in three levels. These comparisons compose a sample size which is ten for different levels. Via comparing the results, the percentage of agreement is calculated between the two subjects in the same level and arranged into a list which shown in the next chapter of the thesis.

# 5. Evaluation and Analysis Experiment Data

## 5.1 Introduction

In previous chapter, in order to understand the design process, the actions in the design protocol are defined level by level via refined EBD coding scheme. In this chapter, the author deals with experiment data. Furthermore, the statistic tool – hypothesis test is utilized in this thesis. In general, the level of details is defined as the deepness of understanding of the design process. Usually, it depends on experience, knowledge and the tools that used to analyze the design process. Uncertainty means the lack of certainty. Limited knowledge makes it impossible to describe exactly outcome or more than one possible result. The purpose of hypothesis test is aim at making sure the existed relation between the level of details and uncertainty.

## 5.2 Processing protocol data

Five subjects analyze the same design protocol. The verbal protocol is respectively analyzed by three levels of EBD coding scheme. Hence, each level has five results. The author chooses two of five results which from the same level to make a comparison. Therefore, each level is constituted by 10 comparisons. In the study, each ten comparisons from three levels compose a group which called sample 1, sample 2, and sample 3.

For example, the verbal protocol is divided into 112 segmentations. There are three factors in level 1. In the experiment, the coders should remove the option which is impossible to define the action. Two results of the same segmentation are compared. If the three options are same, the percentage of agreement is 100%. If the two options are same, the percentage of agreement is 66.7%. If one option is same, the percentage of agreement is 33.7%. If no option is same, the percentage is obviously 0. The author derives an excel form to record the percentage of agreement and calculate the average of the agreements. At the end, each sample includes ten averages of percentage of agreement. Table 5.1 shows all the results of the design protocol.

**Table 5.1 Comparison of agreement of different levels**

| # | Sample 1 | Sample 2 | Sample 3 |
|---|---|---|---|
| 1 | 0.8452 | 0.8232 | 0.8202 |
| 2 | 0.8720 | 0.8786 | 0.8304 |
| 3 | 0.8929 | 0.7893 | 0.7577 |
| 4 | 0.8274 | 0.7679 | 0.6263 |
| 5 | 0.9137 | 0.8482 | 0.8214 |
| 6 | 0.8244 | 0.7732 | 0.7883 |
| 7 | 0.8542 | 0.6893 | 0.6843 |
| 8 | 0.8452 | 0.7661 | 0.7768 |
| 9 | 0.8363 | 0.7411 | 0.6224 |
| 10 | 0.8601 | 0.7804 | 0.6186 |

The author uses Matlab to calculate the mean and the variance of these three samples.

The program of the Matlab is expressed as follows:

```
N1= [0.8452,0.8720,0.8929,0.8274,0.9137,0.8244,0.8542,0.8452,0.8363,0.8601];

N2= [0.8232,0.8786,0.7893,0.7679,0.8482,0.7732,0.6893,0.7661,0.7411,0.7804];

N3= [0.8202,0.8304,0.7577,0.6263,0.8214,0.7883,0.6843,0.7768,0.6224,0.6186];

mu1 = mean(N1);

mu2 = mean(N2);

mu3= mean(N3);

v1 = var(N1);

v2 = var(N2);
```

v3 = var(N3);

Then we can get the results:

$\overline{x_1}$= 85.71%, $s_1$=8.21, $n_1$=10;

$\overline{x_2}$= 78.57%, $s_2$=29.03, $n_2$= 10;

$\overline{x_3}$= 73.46%, $s_3$=77.29, $n_3$= 10;

## 5.3 Hypothesis test

As mentioned before, hypothesis test is used to identify the relations between level of details and uncertainty in the design process. Therefore, there are three sub-hypotheses which prove the relationships between different levels. In this study, T-test is used in the hypothesis test. The reasons are explained as follows:

(1). The size of the sample is too small. There are just 10 data in each sample.

(2). The normal distributions have different means, and the variances are unknown.

In hypothesis test, when the null hypothesis is rejected, the type 1 has occurred. The probability is denoted as α. In general procedure of hypothesis test, the value of α is specified. Hence, in this test, I specified α=0.1.

### 5.3.1 Uncertainty between the first and second level

Sample 1 belongs to the distribution of level 1 and sample 2 belongs to distribution level 2. Using Matlab, the mean and variance of sample are calculated, but the mean and variance of population are unknown.

#### 5.3.1.1 Comparing the mean of the first and second level

$\mu_1$ is the mean of distribution 1, and $\mu_2$ is the mean of distribution 2.

$s_1$ is the variance of distribution 1, and $s_2$ is the variance of distribution 2.

$\Delta_0$ is a specified value, in this case, I specify $\Delta_0$=0.

$H_0$ represents that the mean of distribution 1 and 2 are the same.

$H_1$ expresses that the mean of distribution 1 is larger than distribution 2.

Hence, the hypotheses are:

$$H_0: \mu_1 - \mu_2 = \Delta_0$$

$$H_1: \mu_1 - \mu_2 > \Delta_0$$

$$\text{Test statistic: } t_0^* = \frac{\overline{x_1} - \overline{x_2} - \Delta_0}{\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}} = 2.37$$

$$v = \frac{\left(\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}\right)^2}{\frac{\left({s_1}^2/n_1\right)^2}{n_1+1} + \frac{\left({s_2}^2/n_2\right)^2}{n_2+1}} - 2 = 10.75 \approx 11$$

$$t_{0.10/2,11} = 1.796$$

If $H_0: \mu_1 - \mu_2 = \Delta_0$ is true, then the $t_0^* = \frac{\overline{x_1} - \overline{x_2} - \Delta_0}{\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}}$ is distributed approximately as t

with degrees of freedom given by Test statistic: $t_{0.05/2,11}$= 1.796

According to the results, $H_0$ is rejected.

Hence, the mean of the first level is larger than the second level.

#### 5.3.1.2 Comparing the variance between the first and second level

In this thesis, the samples are from different distributions, and the variances of different distributions are unknown. Thus, I assume that $\sigma$ is the variance of the distribution.

$\sigma_1$ is the variance of distribution 1, and $\sigma_2$ is the variance of distribution 2.

$H_0$ represents that the variance of distribution 1 and 2 are the same.

$H_1$ expresses that the variance of distribution 2 is larger than distribution 1.

Hence, hypotheses are:

$$H_0: \sigma_1{}^2 = \sigma_2{}^2$$

$$H_1: \sigma_2{}^2 > \sigma_1{}^2$$

$$F_0 = \frac{\text{between} - \text{group variability}}{\text{within} - \text{group variability}}$$

$F_{\alpha,n_1-1,n_2-1}$=2.44

Step 1: calculate the mean within each group

$\overline{Y_1}$=85.71%

$\overline{Y_2}$=78.57%

Step 2: calculate the overall mean

$\overline{Y} = \frac{\overline{Y_1}+\overline{Y_2}}{2} = \frac{85.71\%+78.57\%}{2}$=82.14%

Step 3: calculate the "between-group" sum of squares:

$S_B = n * (\overline{Y_1} - \overline{Y})^2 + n * (\overline{Y_2} - \overline{Y})^2$=0.026

The between-group degrees of freedom is one less than the number of groups

$f_b$=2-1=1

$MS_B$=0.026/1=0.026

Step 4: calculate the "with-in group" sum of squares. Begin by centering the data in each group. The results are listed in Table 5.2.

The within-group sum of squares in the sum of squares of all 20 values in this table

$S_W$=(0.00014+0.00022+0.00128+0.00088+0.00320+0.00107+0.00001+0.00014+0.00043 +0.00001)+(0.0014+0.0086+0.00001+0.00392+0.00015+0.00929+0.00038+0.00199+0.0 0003)=0.00864+0.02577=0.034

$f_W$= 2*(10-1)=18

$MS_W$=0.034/18=0.002

**Table 5.2 Centering the date in the first and second level**

| # | Sample 1 | Sample 2 |
|---|---|---|
| 1 | -0.0119 | 0.0375 |
| 2 | 0.0149 | 0.0929 |
| 3 | 0.0358 | 0.0036 |
| 4 | 0.0297 | 0.00001 |
| 5 | 0.0566 | 0.0626 |
| 6 | -0.0327 | 0.0125 |
| 7 | -0.0029 | 0.0964 |
| 8 | -0.0119 | 0.0196 |
| 9 | -0.0208 | -0.0449 |
| 10 | 0.003 | -0.0053 |

$F_0 = \frac{MS_B}{MS_W}$=13> $F_{\alpha,n_1-1,n_2-1}$=2.44

Therefore, $H_0$ is rejected. The variance of the first level is smaller than the second level.

Confidence interval on the difference on means, variances unknown:

If $\overline{x_1}$, $\overline{x_2}$, ${s_1}^2$ and ${s_2}^2$ are the means and variances of sample 1 and 2. The sizes of the two samples are $n_1$ and $n_2$ respectively. These samples are from two independent normal population with unknown and unequal variances, then an approximate 100(1-α)% confidence interval on the difference in means $\mu_1$-$\mu_2$ is:

$$\overline{x_1} - \overline{x_2} - t_{\alpha/2,v}\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}} \le \mu_1 - \mu_2 \le \overline{x_1} - \overline{x_2} + t_{\alpha/2,v}\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}$$

$$3.67\% \le \mu_1 - \mu_2 \le 10.61\%$$

### 5.3.2 Uncertainty between the second and third level

#### 5.3.2.1 Comparing the mean of the second and third level

$\mu_3$ is the mean of sample 3, and $\mu_2$ is the mean of sample 2.

$s_3$ is the variance of sample 3, and $s_2$ is the variance of sample 2.

$\Delta_0$ is a specified value, in this case, I specify $\Delta_0$=0.

$H_0$ represents that the mean of distribution 2 and 3 are the same.

$H_1$ expresses that the mean of distribution 2 is larger than distribution 3.

Hence, hypotheses are:

$$H_0: \mu_2 - \mu_3 = \Delta_0$$

$$H_1: \mu_2 - \mu_3 > \Delta_0$$

$$\text{Test statistic: } {t_0}^* = \frac{\overline{x_2} - \overline{x_3} - \Delta_0}{\sqrt{\frac{{s_3}^2}{n_3} + \frac{{s_2}^2}{n_2}}} = 1.57$$

$$v = \frac{(\frac{s_3^2}{n_3}+\frac{s_2^2}{n_2})^2}{\frac{({s_3}^2/n_3)^2}{n_3+1}+\frac{({s_2}^2/n_2)^2}{n_2+1}} - 2 = 18.26 \approx 19$$

$$\alpha = 0.10$$

$$t_{0.10/2,19} = 1.729$$

If $H_0: \mu_2 - \mu_3 = \Delta_0$ is ture, then the ${t_0}^*$ is distributed approximately as t with degrees of freedom given by Test statistic: $t_{0.10/2,19} = 1.729$

According to the results, the $H_0$ is rejected.

Hence, the mean of the second level is larger r than the third level.

#### 5.3.2.2 Comparing the variance between the second and third level

$\sigma_3$ is the variance of distribution 3, and $\sigma_2$ is the variance of distribution 2.

$H_0$ represents that the variance of distribution 3 and 2 are the same.

$H_1$ expresses that the variance of distribution 3 is larger than distribution 2.

Hence, hypotheses are:

$H_0: {\sigma_2}^2 = {\sigma_3}^2$

$$H_1: {\sigma_3}^2 > {\sigma_2}^2$$

$$F_0 = \frac{\text{between} - \text{group variability}}{\text{within} - \text{group variability}}$$

$F_{\alpha,n_1-1,n_2-1} = 2.44$

**Step 1**: calculate the mean within each group

$\overline{Y_2}$=78.57%

$\overline{Y_3}$=73.46%

**Step 2**: calculate the overall mean

$\overline{Y} = \frac{\overline{Y_3}+\overline{Y_2}}{2} = \frac{73.46\%+78.57\%}{2}$=76.02%

**Step 3**: calculate the "between-group" sum of squares:

$S_B = n * (\overline{Y_2} - \overline{Y})^2 + n * (\overline{Y_3} - \overline{Y})^2$=0.02

The between-group degrees of freedom is one less than the number of groups

$f_b$=2-1=1

$MS_B$=0.015/1=0.02

**Step 4**: calculate the "with-in group" sum of squares. Begin by centering the data in each group, the results are shown in Table 5.3.

The within-group sum of squares in the sum of squares of all 20 values in this table

$S_W$=(0.00396+0.01399+0.00084+0.00006+0.00773+0.00017+0.00504+0.00003+0.00037 +0.00036)+(0.00359+0.00491+0.00001+0.01796+0.00373+0.00078+0.00578+0.00027+ 0.01902+0.02008)=0.018+0.076=0.09

**Table 5.3 Centering the data in the second and third level**

| # | Sample 2 | Sample 3 |
|---|---|---|
| 1 | 0.8232-0.7603=0.0629 | 0.8202-0.7603=0.0599 |
| 2 | 0.8786-0.7603=0.1183 | 0.8304-0.7603=0.0701 |
| 3 | 0.7893-0.7603=0.029 | 0.7577-0.7603=-0.0026 |
| 4 | 0.7679-0.7603=0.0076 | 0.6263-0.7603=-0.134 |
| 5 | 0.8482-0.7603=0.0879 | 0.8214-0.7603=0.0611 |
| 6 | 0.7732-0.7603=0.0129 | 0.7883-0.7603=0.028 |
| 7 | 0.6893-0.7603=-0.071 | 0.6843-0.7603=-0.076 |
| 8 | 0.7661-0.7603=0.0058 | 0.7768-0.7603=0.0165 |
| 9 | 0.7411-0.7603=-0.0192 | 0.6224-0.7603=-0.1379 |
| 10 | 0.7804-0.7603=0.0201 | 0.6186-0.7603=-0.1417 |

$f_W$= 2*(10-1) =18

$MS_W$=0.09/18=0.005

$F_0 = \frac{MS_B}{MS_W}$=4> $F_{\alpha,n_1-1,n_2-1}$

Therefore, $H_0$ is rejected. The variance of the second level is smaller the third level.

Confidence interval on the difference on means, variances unknown:

If $\overline{x_2}$, $\overline{x_3}$, ${s_2}^2$ and ${s_3}^2$ are the means and variances of sample 2 and 3. The sizes of the two samples are $n_2$ and $n_3$ respectively. These samples are from two independent normal population with unknown and unequal variances, then an approximate 100(1-α)% confidence interval on the difference in means $\mu_2$-$\mu_3$ is:

$$\overline{x_2} - \overline{x_3} - t_{\alpha/2,v}\sqrt{\frac{{s_2}^2}{n_2} + \frac{{s_3}^2}{n_3}} \le \mu_2 - \mu_3 \le \overline{x_2} - \overline{x_3} + t_{\alpha/2,v}\sqrt{\frac{{s_2}^2}{n_2} + \frac{{s_3}^2}{n_3}}$$

$$0.46\% \le \mu_2 - \mu_3 \le 11.53\%$$

### 5.3.3 Uncertainty between the first and third level

#### 5.3.3.1 Comparing the mean of the first and third level

$\mu_1$ is the mean of sample 1, and $\mu_3$ is the mean of sample 3.

$s_1$ is the variance of sample 1, and $s_3$ is the variance of sample 3.

$\Delta_0$ is a specified value, in this case, I specify $\Delta_0$=0.

$H_0$ represents that the mean of distribution 1 and 3 are the same.

$H_1$ expresses that the mean of distribution 1 is larger than distribution 3.

Hence, hypotheses are:

$$H_0: \mu_1 - \mu_3 = \Delta_0$$

$$H_1: \mu_1 - \mu_3 > \Delta_0$$

$$\text{Test statistic: } {t_0}^* = \frac{\overline{x_1} - \overline{x_3} - \Delta_0}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_3}^2}{n_3}}} = 4.20$$

$$v = \frac{\left(\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}\right)^2}{\frac{({s_1}^2/n_1)^2}{n_1+1} + \frac{({s_2}^2/n_2)^2}{n_2+1}} - 2 = 13.31 \approx 14$$

$$\alpha = 0.10$$

$$t_{0.10/2,14}= 1.761$$

If $H_0: \mu_1 - \mu_3 = \Delta_0$ is ture, then $t_0{}^*$ is distributed approximately as t with degrees of freedom given by Test statistic: $t_{0.10/2,14}= 1.761$

According to the results, the $H_0$ is rejected,

Therefore, the mean of the first level is larger than the third level.

#### 5.3.3.2 Comparing the variance between the first and third level

$\sigma_1$ is the variance of distribution 1, and $\sigma_3$ is the variance of distribution 3.

$H_0$ represents that the variance of distribution 1 and 3 are the same.

$H_1$ expresses that the variance of distribution 3 is larger than distribution 1.

Hence, hypotheses are:

$$H_0: \sigma_1{}^2 = \sigma_3{}^2$$

$$H_1: \sigma_3{}^2 > \sigma_1{}^2$$

$$F_0=\frac{\text{between−group variability}}{\text{within−group variability}}$$

$$F_{\alpha,n_1-1,n_2-1}=2.44$$

Step 1: calculate the mean within each group

$\overline{x_1}= 85.71\%$, $s_1=8.21$, $n_1=10$;

$\overline{x_2}= 78.57\%$, $s_2=29.03$, $n_2= 10$;

$\overline{x_3}= 73.46\%$, $s_3=77.29$, $n_3= 10$;

$\overline{Y_1}$=85.71%

$\overline{Y_3}$=73.46%

Step 2: calculate the overall mean

$\overline{Y} = \frac{\overline{Y_1}+\overline{Y_2}}{2} = \frac{85.71\%+73.46\%}{2}$=79.59%

Step 3: calculate the "between-group" sum of squares:

$S_B = n * (\overline{Y_1} - \overline{Y})^2 + n * (\overline{Y_3} - \overline{Y})^2$=0.04+0.04=0.08

The between-group degrees of freedom is one less than the number of groups

$f_b$=2-1=1

$MS_B$=0.08/1=0.08

Step 4: calculate the "with-in group" sum of squares. Begin by centering the data in each group. The results are represented in Table 5.4.

**Table 5.4 Centering the date in the first and third level**

| # | Sample 1 | Sample 3 |
|---|---|---|
| 1 | 0.8452-0.7959=0.0493 | 0.8202-0.7959=0.0243 |
| 2 | 0.8720-0.7959=0.0761 | 0.8304-0.7959=0.0345 |
| 3 | 0.8929-0.7959=0.097 | 0.7577-0.7959=-0.0382 |
| 4 | 0.8274-0.7959=0.0315 | 0.6263-0.7959=-0.1696 |
| 5 | 0.9137-0.7959=0.1178 | 0.8214-0.7959=0.0255 |
| 6 | 0.8244-0.7959=0.0285 | 0.7883-0.7959=-0.0076 |
| 7 | 0.8542-0.7959=0.0583 | 0.6843-0.7959=-0.1116 |
| 8 | 0.8452-0.7959=0.0493 | 0.7768-0.7959=-0.0191 |
| 9 | 0.8363-0.7959=0.0404 | 0.6224-0.7959=-0.1733 |
| 10 | 0.8601-0.7959=0.0642 | 0.6186-0.7959=-0.1773 |

The within-group sum of squares in the sum of squares of all 20 values in this table

$S_W$=(0.00243+0.00579+0.00941+0.00099+0.01387+0.00081+0.00340+0.00243+0.00163 +0.00412)+(0.107)=0.15

$f_W$= 2*(10-1)=18

$MS_W$=0.15/18=0.008

$F_0 = \frac{MS_B}{MS_W}$=10> $F_{\alpha,n_1-1,n_2-1}$

Therefore, $H_0$ is rejected. The variance of level 1 is smaller than that of level 3.

$$\overline{x_1} - \overline{x_3} - t_{\alpha/2,v}\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_3{}^2}{n_3}} \le \mu_1 - \mu_3 \le \overline{x_1} - \overline{x_3} + t_{\alpha/2,v}\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_3{}^2}{n_3}}$$

$$7.1\% \le \mu_1 - \mu_3 \le 17.40\%$$

According to the hypothesis test, several results are described as follows:

- Using the EBD coding scheme, the coders can get high percentage of agreement in the same protocol.
- The lower level can get a higher percentage of agreement and smaller variance. However, the information of the cognitive process is insufficient.
- The higher level can further refine the actions, even more, gain more information of the design process. Simultaneously, the more refined action will decrease the percentage of the agreement and increase the variance as well. That is the limitation of refined EBD coding scheme.

# 6. Conclusion and Future Work

## 6.1 Conclusion

The objective of this study is refining the EBD coding scheme and applying the scheme into a case study and at last determining the level of details in design research. EBD scheme codes the designers' actions into several levels. In the present thesis, the coding scheme is developed of three levels. The first level includes three categories: collecting, processing and expressing. The second level has five basic parts – identifying problem, searching information, generating solution, expressing solution and evaluating solution. The third level is formed by 7 parameters: analyzing problem, identifying conflicts or (new) requirements, searching synthesis knowledge, searching evaluation knowledge, generating solution, expressing solution and evaluating solution. In the third level, the author focuses on refining the noun word "information".

In the study, we not only refine EBD coding scheme, but also apply the scheme in a design protocol. Based on the experiment results, the author can prove that the design process is completed described by the EBD coding scheme. Via hypothesis tests, some conclusions are shown as follows:

- Applying EBD coding schemes, coders can define mostly actions in the design process and gain a high percentage of agreement for the definition in the same protocol data.
- The first level achieves a higher percentage of agreement and a lower variance.
- The third level acquires a lower percentage of agreement and a higher variance.

According to data analysis, although the first level could get a high agreement and low variance, it does not mean that first level is better. It could not grasp adequate information from the design protocol. However, the third level could get a large amount of information, but the agreement is low, and it has a high variance. Hence, the relationship can be express that uncertainty increases when the level of details is deeper under the fixed technique.

## 6.2 Future work

Based on the results of the hypothesis test, we can find the relations between level of details and uncertainties in the design research. In the future, we focus on developing deeper level of coding scheme which helps grasp more information related to "thinking" process. In current situation, the deeper the subjects define the activities of design process, the more uncertainties arise. This challenge exists because the activities of the brain cannot be measured directly. In order to solve this problem, we need to have the aid of other techniques, such as electroencephalogram (EEG) and heart rate variability (HRV) methodology, to reduce the fuzzy parts of segmenting.

EEG is a methodology which detects the activity by using electrodes placed on the scalp(Prior, 1984). As we known, different positions of the brain control different function or activity. When the designer solves the problem, electrodes are placed over the different parts of the brain, such as frontal, the parietal, the occipital and the temporal lobes of the brain.

By means of EEG, the change of the activity is distinctly detected. The dynamic of the brain contributes to discriminate among the fuzzy actions.

During the design process, the metal stresses can be indicated by movements of the participants. In the Yerkes-Dodson law, the correlation between designer performance and metal stress can be described as a U-shaped curve (Yerkes & Dodson, 1908).

HRV is a methodology which describes the variance of the consecutive heartbeats. Consequently, variations of the mental stress get rise of the change of the frequency of the heartbeat. The similar frequency of heartbeat conducts to identify the cognitive activity.

In the meanwhile, the coding scheme needs to develop more levels to deeply describe the design protocol and gain more details of the design process. It helps designers to quantify and improve the design process.

In the future, more technologies and equipment will be applied for the developing coding scheme to distinguish the cognitive actions.

## Reference


Anders, K., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251. doi:10.1037/0033-295X.87.3.215

Andy Dong. (2006). How am I doing? The language of appraisal in design. In J. S. Gero (Ed.), *Design Computing and Cognition ’06* (pp. 385–404). Springer Netherlands. Retrieved from http://www.springerlink.com/content/q2482u0101764341/abstract/

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182. doi:10.1016/0364-0213(89)90002-5

Conati, C., & Vanlehn, K. (2000). Toward Computer-Based Support of Meta-Cognitive Skills: a Computational Framework to Coach Self-Explanation. *International Journal of Artificial Intelligence in Education (IJAIED)*, *11*, 389–415.

Cross, N. (2008). Engineering Design Methods: Strategies for Product Design. John Wiley & Sons.

Dong, A., Kleinsmann, M., & Valkenburg, R. (2009). Affect-in-cognition through the language of appraisals. *Design Studies,* *30*(2), 138–153. doi:10.1016/j.destud.2008.12.003

French, M. J. (1998). *Conceptual Design for Engineers*. Springer.

Gero, J. S., & Kannengiesser, U. (2004). The situated function–behaviour–structure framework. *Design Studies*, *25*(4), 373–391. doi:10.1016/j.destud.2003.10.010

Gero, J. S., & Kannengiesser, U. (2006). A Function-Behaviour -Structure Ontology of Processes. In J. Gero (Ed.), Design Computing and Cognition ’06 (pp. 407–422). Springer Netherlands. Retrieved from http://www.springerlink.com/content/m184672747184748/abstract/

Guilford, J. P. (1988). Some Changes in the Structure-of-Intellect Model. *Educational and Psychological Measurement*, *48*(1), 1–4. doi:10.1177/001316448804800102

Hayes, S. C. (1986). The case of the silent dog—Verbal reports and the analysis of rules: A review of Ericsson and Simon’s Protocol Analysis: Verbal Reports as Data. *Journal of the Experimental Analysis of Behavior*, *45*(3), 351–363. doi:10.1901/jeab.1986.45-351

Herbsleb, J. D., Klein, H., Olson, G. M., Brunner, H., Olson, J. S., & Harding, J. (1995). Object-Oriented Analysis and Design in Software Project Teams. *Human–Computer Interaction*, *10*(2-3), 249–292. doi:10.1080/07370024.1995.9667219

Hernandez, N. V., Shah, J. J., & Smith, S. M. (2010). Understanding design ideation mechanisms through multilevel aligned empirical studies. *Design Studies*, *31*(4), 382–410. doi:10.1016/j.destud.2010.04.001

Howard, S. (1997). Trade-off decision making in user interface design. *Behaviour & Information Technology*, *16*(2), 98–109. doi:10.1080/014492997119941

Jin, Y., & Chusilp, P. (2006). Study of mental iteration in different design situations. *Design Studies*, *27*(1), 25–55. doi:10.1016/j.destud.2005.06.003

Lemut, E., Dettori, G., & Boulay, B. D. (Eds.). (1993). *Cognitive Models and Intelligent Environments for Learning Programming*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Mullins, C. A., Atman, C. J., & Shuman, L. J. (1999). Freshman engineers' performance when solving design problems. *IEEE Transactions on Education*, *42*(4), 281 –287. doi:10.1109/13.804533

Nguyen, T. A., & Zeng, Y. (2012). *A Theoretical Model of Design Creativity: Dynamics and Govering Factors* (academic report). CIISE, Concordia University.

P, J. (1956). The structure of intellect. *Psychological Bulletin*, *53*(4), 267–293. doi:10.1037/h0040755

Pennington, N., Nicolich, R., & Rahm, J. (1995). Transfer of Training Between Cognitive Subskills: Is Knowledge Use Specific? *Cognitive Psychology*, *28*(2), 175–224. doi:10.1006/cogp.1995.1005

Prior, P. F. (1984). Fundamentals of EEG Technology. Vol. 1 Basic Concepts and Methods. *Journal of Neurology, Neurosurgery, and Psychiatry*, *47*(3), 326.

Sim, S. K., & Duffy, A. H. B. (2003). Towards an ontology of generic engineering design activities. *Research in Engineering Design*, *14*(4), 200–223. doi:10.1007/s00163-003-0037-1

Simon, H. A., & Kaplan, C. A. (1989). Foundations of cognitive science. In *Foundations of Cognitive Science* (pp. 1–47). Cambridge, MA, US: The MIT Press.

Spiring, F. (2007). Introduction to Statistical Quality Control. *Technometrics*, *49*(1), 108–109. doi:10.1198/tech.2007.s465

Stempfle, J., & Badke-Schaub, P. (2002). Thinking in design teams - an analysis of team communication. *Design Studies*, *23*(5), 473–496. doi:10.1016/S0142-694X(02)00004-2

Suwa, M., Purcell, T., & Gero, J. (1998). Macroscopic analysis of design processes based on a scheme for coding designers' cognitive actions. *Design Studies*, *19*(4), 455–483. doi:10.1016/S0142-694X(98)00016-7

Wang, J., & DONG, A. (2007). How am I doing 2: Computing the language of appraisal in design. Presented at the International conference on engineering design ICED'O7, Paris, France.

Yong, Z. (2008). Recursive object model (ROM)—Modelling of linguistic information in engineering design. *Computers in Industry*, *59*(6), 612–625. doi:10.1016/j.compind.2008.03.002

Zeng, Y. (2011). Environment-Based Design (EBD). *ASME Conference Proceedings*, *2011*(54860), 237–250. doi:10.1115/DETC2011-48263

Zeng, Y., & P.Gu. (1999). A science-based approach to product deign theory part 1: formulation and formalization of design process (Vol. 15, pp. 331–339). Presented at the Robotics and Computer integrated Manufacturing, Elsevier Science Ltd.