

THE DIFFERENCES BETWEEN RETROSPECTIVE AND CONCURRENT PROTOCOLS IN REVEALING THE PROCESS- ORIENTED ASPECTS OF THE DESIGN PROCESS

JOHN S. GERO AND HSIEN-HUI TANG
Key Centre of Design Computing and Cognition
Department of Architectural and Design Science
University of Sydney NSW 2006 Australia
{john, hhtang}@arch.usyd.edu.au

Abstract. This paper presents the results of studying a single designer using protocol analyses and examines the implications of the results on studies of design thinking. It contrasts two types of protocols: concurrent protocols and retrospective protocols. The results indicate that concurrent and retrospective protocols both produce very similar results in terms of exploring the process-oriented aspects of designing. As a result, it is argued there is no associated interference with the ongoing design process when using concurrent protocols.

Keywords. design cognition, protocol studies, case study

Introduction

There is an increasing interest in understanding how human designers design. Part of this interest comes from a need to be able to develop appropriate computational support tools for designers, and part from a need to provide a basis for models of designing. Protocol analysis [1, 2] has become the prevailing experimental technique for exploring this understanding of designing [3]. Two types of protocol approaches have been developed: concurrent and retrospective [4]. In obtaining concurrent protocol, the subjects are required to design and verbalize thoughts simultaneously, while in retrospective protocols, subjects are asked to design first and then retrospectively report the design processes with or without the visual aids provided by the videotapes documenting their own design processes. Generally, concurrent protocols have been utilized when focusing on the process-oriented aspect of designing, which is largely based on the information processing view proposed by Simon [5]. Retrospective protocols have been utilized when focusing on the cognitive content aspect, being concerned with the notion of reflection in action proposed by Schön [6]. Normally design researchers choose one or other methodology depending on their goals.

This paper describes an experiment to examine the similarities and differences between the results produced by these two approaches in an attempt to distinguish the advantages of each method.

1. Protocol Studies of Human Designers

During concurrent protocols, also called the “think aloud method”, subjects design and simultaneously verbalize their design thoughts. Studies using concurrent protocols reveal details of sequences of information processes reflecting the designer’s short-term memory (STM). It is claimed designers can be involved with concurrent protocols without altering their cognitive processes [1]. However, some researchers argue that thinking aloud interferes with the thinking processes and some aspects of the design process could not be revealed [7].

During retrospective protocols subjects retrieve the trace of the preceding cognitive processes and reveal information preserved partially in STM and partially stored in long-term memory (LTM). Characteristics of human memory may seriously impair the results, so the retrieved data from LTM may have details omitted or may be generated by reasoning rather than recall. As a result, some researchers utilize videotapes of the design session as cues during retrospection to assist in the recall of the design activity [8]. In this study we combine both methods and utilize both protocols with a single subject in an extended experiment. As a consequence, we are in the position to obtain more complete protocols from design sessions than either method alone. If there is a substantial agreement between the results from both protocols, then we can have more confidence in the results.

Gero and Mc Neill [9] have proposed the most complete of the coding schemes to understand the process-oriented aspects of designing. It consists of problem domains and design strategies highly related to design processes. The information categories, proposed by Suwa and Tversky [10] and developed by Suwa, Purcell and Gero [8], were established to understand the content-oriented aspects of design. They used notions proposed by Larkin and Simon [11] to define three subclasses to analyze what designers see and possibly think. In this paper we present only the results produced when using the process-oriented coding scheme.

After the collection of data the raw protocol is divided into small units called “segments”. The purpose of segmentation is to facilitate the analysis process because the encoding is based on a single segment that will belong to one or some of the subclasses of one category in the coding scheme. In recent protocol researches [8, 12], the protocol has been divided along the lines of the designer's intentions and actions instead of verbalization events or syntactic markers [1]. The designer's intention is interpreted to produce each segment, and each segment presents one single intention of the designer in design process.

Goldschmidt [13] proposed a definition of segmentation where the protocol is segmented by the designer's intention in her protocol studies of architects. She divided the design process into “moves” and “arguments”. Moves divide a stream of design activities into the smallest units of design reasoning present, a coherent proposition pertaining to an

entity that is being designed. Arguments are the smallest sensible statements about the design or aspects of it, being related to a particular design move. Generally, one move consists of one or two arguments. The scale of segments proposed by Gero and Mc Neill is more like the scale of arguments, whilst the scale of segments proposed by Suwa, Purcell and Gero is more like the scale of moves.

Although the definition of a segment is precise, it was still vague in its application in terms of how to divide the protocol into appropriate segments in some particular situations. The methods of segmentation in recent papers [8, 12] are similar to Goldschmidt's definition, while the relationship between one segment and the encoding code are different. In the Gero and Mc Neill paper, one encoding code corresponds to one segment, so the length of segments is related to the subcategory. In contrast, in Suwa, Purcell, and Gero's paper, there may be more than one code in one segment, so the subcategory does not affect the segments while designers' intentions do. Moreover, the bases of segmentations are different in the two coding schemes. The coding schemes proposed by Gero and Mc Neill are principally based on the transcripts, while the schemes proposed by Suwa, Purcell and Gero are essentially based on the designer's actions in the video. As a result, the meanings of segmentations are different in these two coding schemes. In this study we use the definition proposed by Gero and Mc Neill, which is concordant with our examination in the process-oriented aspects of design.

2. The Experiment

We recruited third year undergraduate design students from the University of Technology Sydney. The experiment was set up in one of the subjects' studios. The equipment used was two Hi-8 cameras and one microphone with an amplifier.

In order to have both concurrent and retrospective protocols from a single design session, we used four phases: warm-up exercises, think aloud while designing, retrospection and final interview. Two warm-up exercises were used to accustom the subject to the methodology. During the think aloud phase the subject designed and generated a concurrent protocol at the same time. In the retrospective phase, he generated a retrospective protocol with the aid of the videotape of his designing. Generally these four phases took about 4 hours.

We used one subject to produce both the concurrent and retrospective protocols instead of two subjects each producing one type of protocol. We did this because the differences between two individual designers are hard to compensate for. Further, commencing with the concurrent protocol may benefit the retrospective protocol to some extent because the process of the concurrent protocol will reinforce long-term memory that will then assist the retrospective process.

Each type of protocol has two videotapes: one providing a macroscopic view and the other a microscopic view. The former provides

the general gestures of the subject, while the latter captures the detailed movements of sketching. After transcribing the audio part of the videotapes, the raw protocols were encoded by one encoder in this study. There was one ten-day break between the first and the second encoding processes. The discrepancies between them are than encoded again in an arbitrated phase. The results presented here are based on a single subject who had the best quality of both concurrent and retrospective protocols although a number of subjects were involved and examined.

Table 1 shows coding consistency between phases of this subject, indicating that the result of encoding became consistent using this process.

TABLE 1. The coding consistency between different coding phases

	First and second coding	Second and arbitrated coding
Concurrent protocol	55 %	85 %
Retrospective protocol	64 %	85 %
Overall	60 %	85 %

3. Protocol Analysis Results

3.1 OBSERVED DIFFERENCES

The results of the experiments indicate some differences between the two types of protocols.

First, concurrent protocols appear to reveal more information in the beginning of the design process especially when the subjects were trying to perceive the problem. In retrospective protocols subjects could not adequately remember these early processes even with the help of the video.

According to previous studies, this period is an important part of problem finding or problem formulation in which the designers internalized their external thoughts [7] and set the scene for subsequent designing behaviour [14]. However, even the information revealed by the concurrent protocol may not be sufficient to provide insight of this important period. It becomes possible to explore this thought process when subjects externalize the internalized external information. The example is when they wrote down the issues they wanted to pursue in the design after reading through the brief.

Second, during concurrent protocols sometimes subjects paused their speech. These pauses were regarded as the transition of attention or non-verbal thought in previous studies. Little information about the pause was provided by the concurrent protocol. However, we can find some information about these pauses in the retrospective protocol because subjects sometimes can recall the thinking process in this period. The

periods of pause often happened when the designer drew or examined his sketches intensively.

To obtain a detailed comparison, we carried out a statistical analysis of the results of each type of encoded protocol.

3.2 QUANTITATIVE DIFFERENCES

In all our experiments the lengths of retrospective protocols are longer than concurrent protocols, but this does not directly indicate that retrospective protocols could reveal more information than concurrent protocols. The situation may result from subjects having more time to talk in retrospective protocols or from the effects of simultaneous sketching in concurrent protocols. Consequently, we used the segments to assess the amount of information in both types of protocols. In this study the segments were divided by designers' intention according to the definition given in Gero and Mc Neill [12].

After segmentation, the number of segments in the retrospective protocol is higher than in the concurrent protocol, Table 2. The concurrent protocol contains 218 segments while the retrospective protocol contains 276 segments, i.e. 26 percent more. Since the segments represent the intention of designers, according to the definition given earlier, we argue that retrospective protocols reveal more intentional information in this data set. Moreover, the average length of one segment of concurrent protocol is longer than that of retrospective protocol. This is concordant with the general argument that concurrent verbalization affects the design process and vice versa. However, as will be discussed later this alone does not provide evidentiary support as to whether the concurrent verbalization interferes with the retrieval of the content of the design process.

TABLE 2. Descriptive statistics for duration of segment in two types of protocol (unit: second)

	Mean Segment Time	Std.Dev	Std. Error	Maximum Segment Time	Segment Count
Concurrent protocol	7.6	6.4	.44	37	218
Retrospective protocol	6.2	4.3	.26	24	276
Both Protocols	6.8	5.4	.24	37	494

In order to obtain more detailed information about the duration of the design activities, the spectrum of segment lengths is plotted to understand how fast designers change their intentions and the abilities of both protocols to reveal the change of intentions, Figure 1. The spectra of

segment lengths from both protocols are very similar to each other although the segment in retrospective protocol is generally shorter than that in concurrent protocol.

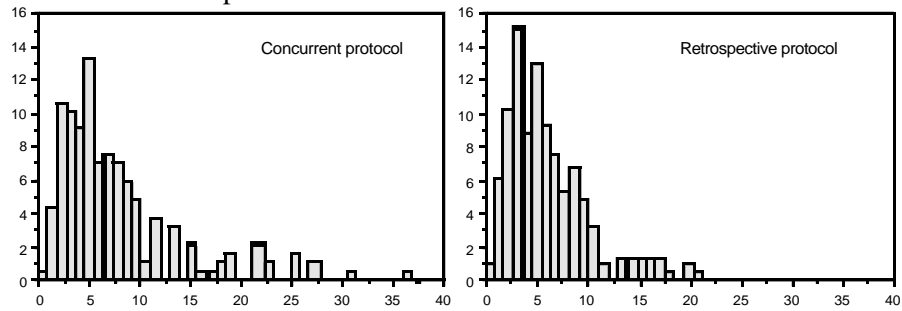


Figure 1. Spectrum of segment lengths in the concurrent protocol (left), and the retrospective protocol (right)

The four most frequent segment lengths for both of them range from 3 to 6 seconds and the means of the segment lengths for the concurrent and retrospective protocols are 7.5 seconds and 6.3 seconds, respectively. The maximums of both protocols are 37 and 24 seconds. The results show that both the concurrent and retrospective are very similar in measuring the frequency of designers' change of intention. In addition, the concurrent protocol has some segments that are longer than those in retrospective protocol. To determine how similar the duration of segments is in both protocols, a comparison of the percentile plot and the Mann-Whitney U test were made, Figure 2. The Mann-Whitney U test shows they are similar to a great extent, while the comparison of the percentiles clearly shows that up to 85% of them are almost the same in length. In the remaining 15%, the duration of segments of concurrent protocol is longer than those in retrospective protocol.

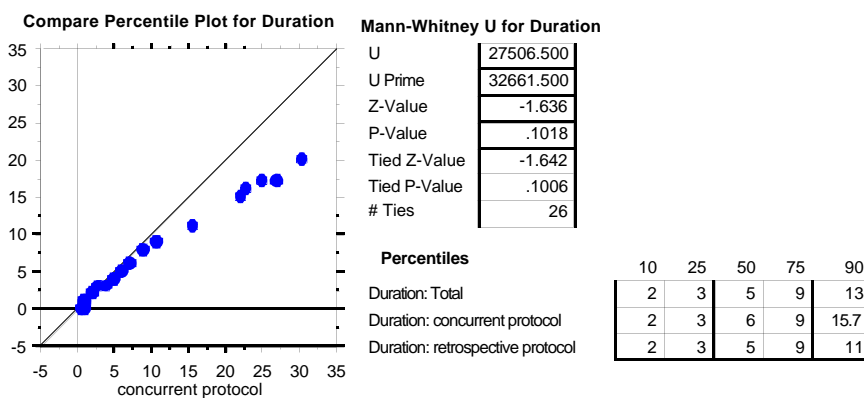


Figure 2. Comparison percentile plot for duration of two types of protocol

This situation is different from the general expectation that retrospective protocols may be longer than concurrent protocols because

there is more time to report. As a result, we can argue that they are similar quantitatively in terms of segments.

3.3 QUALITATIVE DIFFERENCES

To understand the qualitative differences between the two types of protocols, this study compares both the encoded protocols using the process-oriented coding scheme proposed by Gero and McNeill [12]. What encoded protocols represent is the content of the target design process, i.e. how designers navigate the problem domain and what kind of strategies they apply when designing.

After encoding we measured the distribution of each subcategory in Gero and McNeill's coding scheme and statistically examined the differences. Both concurrent and retrospective protocols were from a single design session, so if they showed salient differences in the distributional figures then we could argue that these different types of protocols could reveal different aspects of designing.

The time spent on the four categories of Gero and McNeill's coding scheme are respectively summed for the whole design episode and presented as percentages. Figures 3, 4 and 5 respectively present the analysis by level of abstraction, prototype and macro strategy. Every category is described using Chi-square test values, expected values, and two histograms, the left one for the concurrent protocol and right one for the retrospective protocol.

Gero and McNeill's paper presents the designer's navigation through the problem domain in two orthogonal dimensions, *Level of abstraction* describes the shift of the designer's consideration from high level overall views to low level detailed views. Here, Figure 3, the segment encoded as 0 is the segment when the designer is considering the system as a whole, while the segment encoded as 2 is the segment when the designer is considering parts of the system.

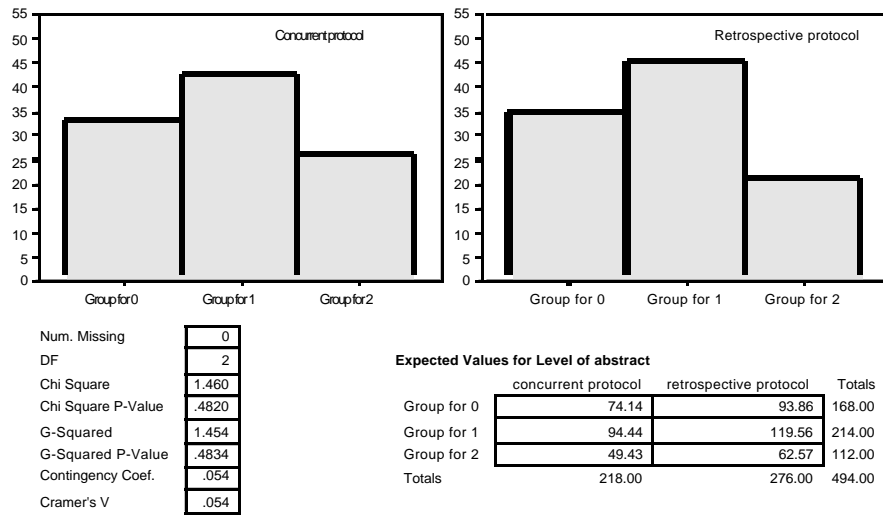


Figure 3. Histograms, Chi-square values, and expected values of level of abstraction in the concurrent protocol (left) and the retrospective protocol (right)

Another dimension to describe the navigation of the problem domain is through function, behavior and structure [9, 15]. *Function* relates to the purpose of an artifact, *Behaviour* relates to the actions or processes of an object or artifact, and reasoning in *Structure* involves the manipulation of objects or their relations to bring about a physical solution. In Figure 4 B stands for behaviour, F for function, and S for structure.

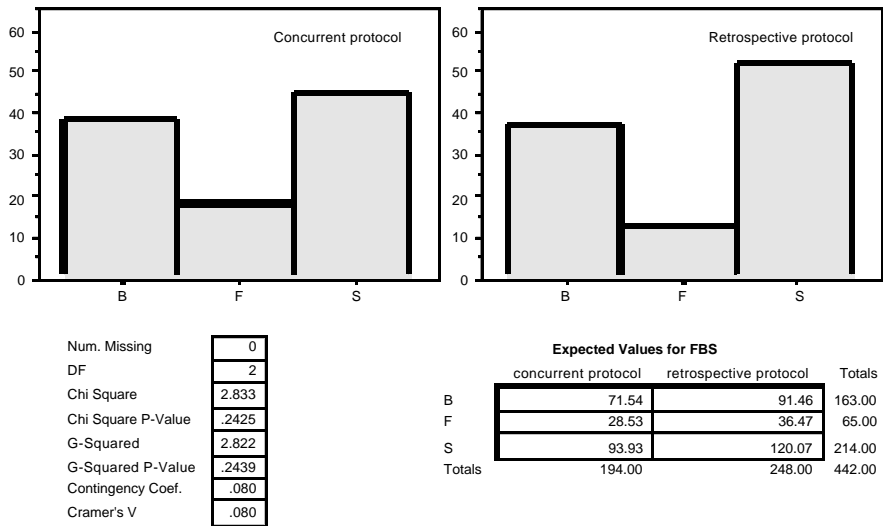


Figure 4. Histograms, Chi-square values, and expected values of prototype in the concurrent protocol (left) and the retrospective protocol (right)

In addition to the dimensions of the navigation of the design problem, there is another means of describing the design process by considering

the design process as a sequence of repeated actions. These repertoires of actions can be identified as being either short term lasting for a few seconds to tens of seconds, or longer term lasting for one minute to a few minutes. The former are called micro strategies, while the latter are called macro strategies in Gero and Mc Neill's paper.

Five distinct macro strategies are identified in the encoding process, Figure 5. In *Bottom up* (Bu) mode, designers examine several alternative structures and their behaviours to modify the requirements. In *Top down* (Td) mode, designers elaborate the functions and related behaviours, and generate sub-goals that need to be addressed in structure. *Decomposing the problem* (De) mode involves the decomposition of either the overall goals or the potential system prior to Top Down design. *Backtracking* (Bt) occurs when the designer is not achieving what has been expected and returns to an earlier process. As a consequence, he goes back over existing work to solve the problem. The *Opportunistic* (Op) strategy occurs when there is an external influence that makes a change of direction advantageous.

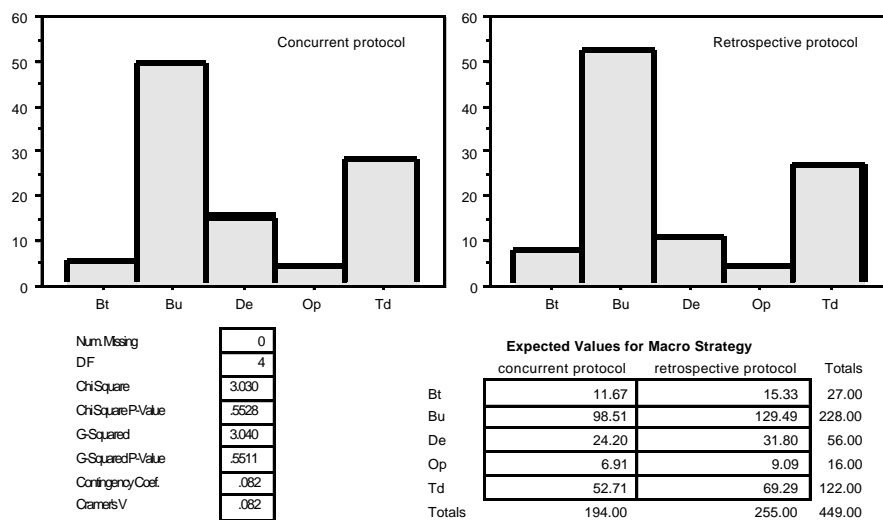


Figure 5. Histograms, Chi-square values, and expected values of macro strategies in the concurrent protocol (left) and the retrospective protocol (right)

We can see that the distributions of concurrent and retrospective protocols are graphically very similar in all of the three categories, Figure 5. Statistically, the low Chi-square values and high p-values indicate that there is little differences between the distributions of the two types of protocols in level of abstraction, prototype, and macro strategy. Although we only have two data sets from a single subject, the expected values demonstrate the validity of the statistical tests in comparing them.

The micro strategy is another other way to categorize the repertoire of the short term design activities. Table 3 presents the different sub-categories we identified in the encoding process. They could be further categorized into analysis, synthesis and evaluation.

TABLE 3. The definitions of categories in micro strategy (after Gero and McNeill (1998))

Analysis	Ap	Analyzing the design problem
	Cp	Clarifying the design problem
	Ep	Evaluating the design problem
Synthesis	Ps	Proposing a solution
	Cl	Clarifying a solution
	Re	Retracting a previous solution
	Dd	Making a design decision
	Co	Consulting external information
	Pp	Postponing a design action
	La	Looking ahead
	Lb	Looking back
Evaluation	An	Analyzing a proposed solution
	Ju	Justifying a proposed solution
	Ca	Calculating on a proposed solution
	Pa	Postponing an analysis of action
	Ev	Evaluating a proposed solution
Explicit Strategies	Ka	Referring to application knowledge
	Kd	Referring to domain knowledge
	Ds	Referring to design strategy
	X	Other

We find some differences between the graphical distributions of these categories in the concurrent and retrospective protocols, Figure 6. Two interesting phenomena are noticeable. First, analyzing the problem, consulting external information and evaluation of the design problem occupy a slightly larger proportion of the activities in the concurrent protocol than in the retrospective. Second, the retrospective protocol occupies a larger proportion than concurrent in evaluating a proposed solution, retracting a previous solution, and proposing a solution. However, due to the small size of the data set, it is not valid to utilize the Chi-square test to examine the degree of the differences so that we cannot state that there is any statistically significant difference

The differences between distributions imply that this concurrent protocol reveals more information related to the functional aspect of the design process where the functions of the design are generated, namely the problem formulation. Comparatively, this retrospective protocol reveals more information in producing solutions and evaluation. These

are related to the designer’s intensive sketching and revising existing depictions.

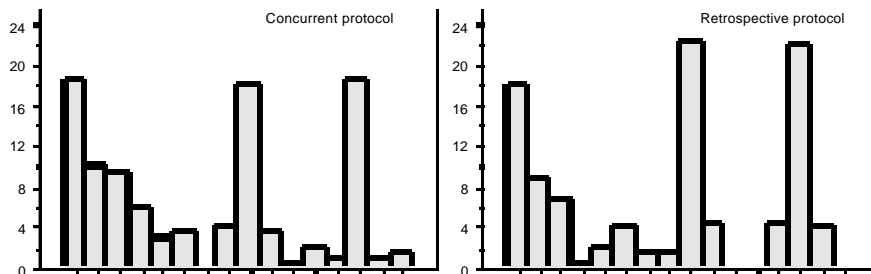


Figure 6. Histograms of micro strategies in the concurrent protocol (left) and the retrospective protocol (right)

Moreover, we compare the distributions of three categories of micro strategies, analysis, synthesis and evaluation from a macroscopic view; this produces a large data set, Figure 7. Statistically, the Chi-square test shows they are less similar to each other, compared to the similarities in levels of abstraction, prototype and macro strategy.

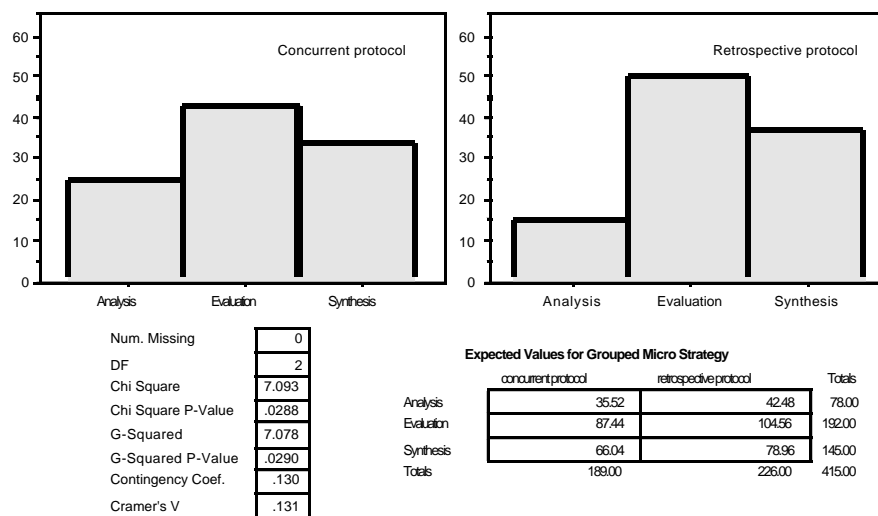


Figure 7. Chi-square values, and expected values of analysis, synthesis and evaluation in concurrent protocol (left) and retrospective protocol (right)

The statistical analyses of level of abstraction, function-behavior-structure, and macro strategies show similarities in revealing the process-oriented aspects of the design process. Although the details vary, the comparison of analysis-synthesis-evaluation in protocols shows the similarity between protocols in the micro strategies level.

4. Conclusions

This study demonstrates some characteristics of concurrent and retrospective protocols by comparing them using a process-oriented coding scheme. Both protocols in our study have the same duration, but the number of segments in the retrospective protocol is greater than in the concurrent. One explanation for the reason this may be that simultaneous sketching and revising influences the reporting during the concurrent protocol. The other explanation is the retrospective protocol is influenced by a rehearsed memory about the design process in the concurrent protocol.

The study by Berry and Broadbent [16] supports these two explanations. They proposed that concurrent verbalization alone has no effect on task performance and reported knowledge of the task. Further, they suggested that the combination of the instruction and concurrent verbalization improves the performance when giving reported knowledge of the task. The reported study here could be regarded as a similar experiment in complex cognitive activities. This result harmonizes with the argument of protocols proposed by Ericsson and Simon (1993). They proposed that concurrent and retrospective protocols are similar to each other, but that retrospective protocols may have details omitted due to the decay of long-term memory. In our study, visual information generated during the design process is available to aid this, so the concurrent and retrospective protocols are similar to each other.

Comparatively, the contents revealed by concurrent and retrospective protocols are similar in this study. Our analyses present their similarities in level of abstraction, function-behavior-structure, macro strategies, and in the analysis-synthesis-evaluation model. As a result, we can argue that concurrent verbalization can reveal the design process to echo the question proposed by Lloyd, Lawson and Scott [7]. They proposed that deficiencies in concurrent protocols cause incompleteness when they are used to reveal the design process; however, we argue that in terms of the process-oriented aspects of designing, concurrent protocols have the same abilities as retrospective protocols. They are still the most efficient and applicable methods in exploring this aspect of the design process.

Further study is needed to explore the differences between concurrent and retrospective protocols in the content-oriented aspects of the design process. In this study, the utility of the Gero and McNeill's method restricts us from drawing conclusions related to the content of the design process. Sketching/sketches, revising, seeing, and these elements emphasized by Schön [6] play important roles in the design process. In order to understand the design process, it is necessary to explore them, but before this we should examine which types of protocols are suitable for this purpose and their differences.

Acknowledgments

This research is supported by a grant from the Australian Research Council, an OPRS scholarship from Australia and a UPA scholarship from the University of Sydney. Computing resources are provided by the Key Centre of Design Computing and Cognition.

References

- [1] Ericsson KA and Simon HA *Protocol Analysis: Verbal Reports as Data* MIT Press, Cambridge, MA (1993)
- [2] van Someren MW, Barnard YF and Sandberg C *The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes*. Academic Press, London (1994)
- [3] Cross N, Christiaans H and Dorst K *Analysing Design Activity* John Wiley, Chichester, UK (1996)
- [4] Dorst K and Dijkhuis J Comparing paradigms for describing design activity *Design Studies* Vol 16 (1995) pp261-274
- [5] Simon HA *Sciences of the Artificial* Cambridge MIT Press, Cambridge, MA (1992)
- [6] Schön DA *The Reflective Practitioner: How Professionals Think in Action* Arena, London (1995)
- [7] Lloyd P, Lawson B and Scott P Can concurrent verbalization reveal design cognition? *Design Studies* Vol 16 (1995) pp237-259
- [8] Suwa M Purcell T and Gero JS Macroscopic analysis of design processes based on a scheme for coding designers' cognitive actions *Design Studies* Vol 19 (1998) pp455-483
- [9] Gero JS and McNeill T An approach to the analysis of design protocols, *Design Studies* Vol 19 (1998) pp21-61
- [10] Suwa M and Tversky B What do architects and students perceive in their design sketches? A protocol analysis *Design Studies* Vol 18 (1997) pp385-403
- [11] Larkin J and Simon HA Why a diagram is (sometimes) worth ten thousand words? *Cognitive Science* Vol 11 (1987) pp65-99
- [12] McNeill, T, Gero, J S and Warren, J Understanding conceptual electronic design using protocol analysis, *Research in Engineering Design* Vol 10 (1998) pp129-140
- [13] Goldschmidt G The dialectics of sketching *Creativity Research Journal* Vol 4 (1991) pp123-143
- [14] Darke J The primary generator and the design process in WE Rogers and WH Ittelson (eds), *New Directions in Environmental Design Research. Proceedings of EDRA 9*, Washington (1978) pp325-337
- [15] Gero JS, Tham KW and Lee HS Behaviour: A link between function and structure in design in DC Brown, MB Waldron and H Yoshikawa (eds), *Intelligent Computer Aided Design* North-Holland, Amsterdam (1992) pp193-225
- [16] Berry D C and Broadbent D E On the relationship between task performance and associated verbalizable knowledge *The Quarterly Journal of Experimental Psychology* Vol 36A (1984) pp209-231.

This is a copy of the paper: Gero, J. S. and Tang, M. (2001) Differences between retrospective and concurrent protocols in revealing the process-oriented aspects of the design process, *Design Studies* **21** (to appear)