UNEXPECTED DISCOVERIES AND S-INVENTION OF DESIGN REQUIREMENTS: A KEY TO CREATIVE DESIGNS

MASAKI SUWA, JOHN GERO AND TERRY PURCELL
Key Centre of Design Computing
Department of Architectural and Design Science
University of Sydney
NSW 2006 Australia

Abstract. *The creative outcome of a design process hinges on whether or not designers are able to find important aspects of a given problem and thereby invent design issues or requirements during the process. How do they do this? What becomes the impetus for the invention of important issues or requirements? So-called "unexpected discoveries", the acts of attending to visuo-spatial features in sketches which were not intended when they were drawn, are believed to contribute to it. The purpose of the present research is to verify this hypothesis. Analysing the cognitive processes of a practising architect in a design session, we found that in about a half of his entire design process there was bi-directional causality between unexpected discoveries and the invention of issues or requirements; not only did unexpected discoveries become the driving force for invention, but also the occurrence of invention, in turn, tended to cause new unexpected discoveries. This has pedagogical implications.*

1. Introduction

How does a design process proceed? Do designers commence with an analysis of important aspects of the given design problem, then synthesise solutions based on this analysis, and finally evaluate the solutions? The answer is no. Designers do not go through these phases in a sequential order in the entire process, but rather analyse, synthesise and evaluate in a more rapid cycle, almost simultaneously. In many cases, it is only after designers synthesise a solution that they are able to detect and understand important issues and requirements of the given problem. Lawson (1990) called this phenomenon "analysis through synthesis". Further, problem-finding behaviours of this sort are strongly associated with creative outcomes. Getzels and Csikszentmihalyi (1976) presented evidence for this from a longitudinal study of art students.

For the same reason, freehand sketches are indispensable for designers in conceptual design processes. It is not until externalizing on paper the ideas of what they think might be a potential solution and inspecting them that designers are able to find new aspects of the problem and to generate new ideas (Robbins, 1992; Goldschmidt, 1994). Our previous study of architects' sketches presented evidence that their thoughts of functional issues in an architectural design task are situated actions; they are born from the process of drawing on paper and perceiving the visuo-spatial features of depictions (Suwa, Gero and Purcell, 1998). Then, a question arises. What aspects in the acts of drawing sketches and perceiving features in them enable a designer to invent important design issues and requirements of a given problem? By "invent" we do not mean that the issue or requirement has been generated for the first time in history, i.e. historical-invention in Boden's terminology (1990), or that a designer has generated it for the first time in his or her life, i.e. psychological-invention. What is meant in this paper is that a designer has generated the issue or requirement for the first time in the current design task, in a way situated in the design setting. We call this "situated-invention (S-invention)".

Freehand sketches are believed to encourage discoveries of unintended features and consequences (Schon and Wiggins, 1992; Goldschmidt, 1994). Making a depiction on paper forces some organization and specificity in terms of visuo-spatial features (Stenning and Oberlander, 1995), regardless of whether or not the sketcher pays attention to them. For example, a depiction necessarily takes some shape and occupies an area of a certain size on paper, even though these visual features may not be intended by the sketcher. When a sketcher makes a new depiction, intending it to hold a spatial relation to some existing depictions, it will automatically produce spatial relations between the new depiction and other existing depictions which the sketcher does not intend. These implicit visuo-spatial features, in turn, may be discovered in an unexpected way by later inspection.

Unexpected discoveries of this sort are believed to become the strong impetus for the invention of important design requirements of a given problem. To quote Goldschmidt,

> "One reads off the sketch more information than was invested in its making. This becomes possible because when we put down on paper dots, lines and other marks, new combinations and relationships among these elements are created that we could not have anticipated or planned for. We discover them in the sketch as it is being made, and some of these new configurations can potentially provide useful clues" (1994, p.164)

She argues that unexpected discoveries of new combinations and relationships among elements will provide useful clues; clues for what? She explains in the same page as follows;

> " ... It is our belief that the purpose of this early sketching activity is primarily to avail oneself of potentially meaningful clues. If picked up, these clues can be used to form and to inform emerging design concepts." (1994, p.164)

It is obvious that what she means by "to inform emerging design concepts" is equal to what we mean by "to invent important issues and requirements". All in all, her argument here suggests that there might be causality from unexpected discoveries to the S-invention of important issues and requirements. No past research, however, has presented empirical evidence for this. The purpose of this paper is to verify this hypothesis, seeking empirical evidence in the cognitive processes of designers.

2. Our Approach

We have looked into the cognitive processes of a practising architect using the technique of protocol analysis. The protocols of his design session were collected as a retrospective report after the session (Suwa and Tversky, 1997). The design session, which lasted for 45 minutes, was to work on the conceptual design of a museum on a given site in a natural environment in the suburb of a large city. The architect was encouraged to draw sketches on tracing paper. His sketching activities were videotaped. In the report session, he reported on what had been on his mind for each stroke of his pencil during the design session, while watching the videotape.

The S-invention of design issues and requirements appear in protocols as the acts of setting up goals to bring them into reality. We found in his protocols many instances of goals of this kind. For example, in a relatively early phase of his process, he decided that visitors to this museum should experience a cheerful and pleasant feeling even before getting to the main building from the parking lot. As another example, he talked about the necessity of service area for the buildings, such as for the delivery of goods and garbage collections. Functional requirements of this sort, the former being psychological and the latter practical, were not given to him as initial requirements, but rather emerged during the process. We coded from the protocols instances of goals to invent functional issues and requirements. We will describe this in Section 3.3.

Another task in this research is to code instances of unexpected discoveries from verbal protocols. The pioneering researches that have established the concept of unexpected discovery (Goldschmidt, 1994; Schon and Wiggins, 1992), however, did not present an explicit definition of it; they just presented convincing examples of unexpected discoveries from protocols in an episodic way, because the aim was to show that unexpected discovery is one of the benefits that sketching provides designers with. We have recently developed a coding scheme that is suitable for examination of cognitive processes of designers (Suwa, Purcell and Gero, 1998). There, we proposed that unexpected discovery can be defined as a class of perceptual actions. We have used this definition to collect instances from the protocols of our architect. The definition will be presented in Section 3.2.

Based on codings of both goals for S-invention and unexpected discoveries, we have examined whether there is any significant causality between the occurrences of both, and if any, what kinds of causality they are. The results will be reported in Section 4.

3. Coding of Unexpected Discoveries and Goals for Inventions

3.1 CODING SCHEME: OVERVIEW

The main thrust of our coding scheme is that it allows for codings of different modes of designers' cognitive actions from protocols. The definitions and codings of both unexpected discoveries and goals for S-invention are derived from this scheme. First, we will give a brief overview of the scheme.

3.1.1 Segmentation

As many previous protocol analysis methods have done, we divide the entire verbal protocol into small units, that is, segmentation. The method of segmentation we employed is to divide the protocol based on the subject's intention (e.g. Goldschmidt, 1991; Van Someren et. al, 1994; Suwa & Tversky 1997; Gero & McNeill, 1998). For example, Goldschmidt (1991) defined a segment, what she calls a "design move", as "an act of reasoning which presents a coherent proposition pertaining to an entity that is being designed". A change in the subject's intention or in the contents of their thoughts or their actions flags the start of a new segment. Consequently, a single segment sometimes consists of one sentence, and sometimes of many.

3.1.2 Different Modes of Cognitive Actions

For each segment, we code different modes of designers' cognitive actions. There are four modes; physical, perceptual, functional and conceptual. The first category, **physical,** refers to actions of (1) drawing new elements (e.g. lines, dots, circles and so on) on paper, (2) paying attention to previously drawn elements, and (3) tracing over, touching, or copying on another sheet previously drawn elements. These all have direct relevance to physical depictions on paper. We collect instances of physical actions from the videotape of the subject's sketching activities.

The second category, **perceptual**, refers to actions of attending to visuo-spatial features of depicted elements. There are four classes of perceptual actions: (1) visual features of elements, such as shapes, sizes, or textures, (2) spatial relations among elements, such as proximity, remoteness, alignment, intersection, connectedness and so on, (3) organization or comparison among elements, such as grouping, uniformity/similarity, contrast/difference, and (4) implicit spaces that exist in-between depicted elements. We collect instances of perceptual actions from verbal protocols, interpreting the semantic contents of the protocol.

It is hypothesised that perceptual actions have dependencies on physical actions. For example, suppose that an architect mentioned the shape of an element drawn before. Attention to the shape is an instance of perceptual action, and is dependent on the physical action of paying attention to the existence of the previously drawn element. As another example, suppose that an architect arranged a new element near a previously drawn element, considering the proximity between the two. The attention to the proximity is an instance of perceptual action, and is dependent on two physical actions: an action of newly depicting the former element, and an action of paying attention to the existence of the latter element. We code not only instances of cognitive actions but also dependencies among actions.

We will describe functional and conceptual actions later, when we describe the coding of goals.

3.1.3 Index of Whether or not Cognitive Actions are New

A cognitive action, whether belonging to physical, perceptual, functional, or conceptual categories, may occur at a segment for the first time since the architect started working on the design task. Or, it may have occurred at a previous segment and is now being repeated. For each cognitive action coded in a segment, we code this information, called 'index', as well. There are two indices; "new" and "old".

If an action has never happened before in the architect's design process, we code the action as "new". Making a new element on paper is a "new" physical action. If an architect considers a shape of an element for the first time, the attention to the shape is a "new" perceptual action. If an architect notices a spatial relation between two elements for the first time, the attention to the spatial relation is a "new" perceptual action.

On the other hand, if an architect revisits an element which was drawn at a previous segment, or repeats an action done at a previous segment, we code the action as "old". What we mean by "a previous segment" could be the immediately previous one of the current segment, or a segment that happened long before. Tracing over a previously drawn element is an "old" physical action. Paying attention to the existence of a previously drawn element is an "old" physical action, too. If an architect has thought of a shape of an element at a previous segment and now mentions the same shape of the same element again, the attention to the shape is an "old" perceptual action.

3.2 DEFINITION OF UNEXPECTED DISCOVERIES

3.2.1 How Unexpected Discovery was Interpreted in Past Literature

Schon and Wiggins (1992) argued that unexpected discovery is an action of noticing consequences that were not intended by the sketcher when he or she drew it. What they mean by "consequences" include many things. They presented, as an example, the sketches of an architectural student in which she arranged six classrooms for six grades in a series of L-shapes. Schon and Wiggins discussed that, even though her initial intention was just to extend the sizes of classrooms, she happened to discover a couple of spaces in front of the classrooms as well as a spatial relation between these spaces, and also thought that these relations can be used for a certain function. "Spaces", "spatial relations" and "function" were all instances of what they mean by unintended "consequences". The "spaces" and the "spatial relation" are visual information that she directly detected in the appearance of a sketch, while "function" is a non-visual meaning that she associated with the visual information In this sense, Schon and Wiggins' version of unexpected discovery includes the discovery not only of visuo-spatial features in sketches but also of concepts that emerge as a result of interpreting visuo-spatial features.

On the other hand, Goldschmidt's version of unexpected discovery is more simple; it includes only the discovery of visuo-spatial features, as obvious in the quotes shown in Section 1. We prefer restricting our definition of unexpected discovery to this simpler version. The reason is that if we include the discovery of non-visual concepts, it would blur the distinction between unexpected discovery and so-called "re-interpretation", another benefit which sketching provides designers with (Goel, 1995, Goldschmidt, 1991). In order to understand that the emergent spaces in front of the L-shaped classrooms could be used for a certain function, the female student must have had one more step of inference, after noticing the existence of those spaces, to interpret the configuration of the spaces. This inference, which is non-visual per se, should be more associated with the act of "re-interpretation".

3.2.2 Our Definition of Unexpected Discovery

Based on this discussion, we defined unexpected discovery as a class of perceptual actions. An unexpected discovery is a "new" perceptual action that has a dependency on "old" physical action(s). This means that if an architect traces over or pays attention to the existence of previously drawn element(s), and notices visual features of or relations among those elements, we say that the perceptual action is an instance of unexpected discovery.

There are three semantically distinct types of unexpected discovery. Table 1 summarises the three types, and Table 2 shows the kinds of physical actions relevant to the definition of unexpected discovery. One is the discovery of a visual feature such as shape, size or texture of a previously-drawn element. It is defined as a "new" perceptual action that has a dependency on an "old" physical action. If an architect traces over a circular line that was originally a simple indication of an area for a function, e.g. entrance hall in the main building of a museum, and now begins to attend to its circular shape for the first time, this is an instance of unexpected discovery of the first type.

The second is the discovery of a spatial or organisational relation among more than one previously-drawn element. It is defined as a "new" perceptual action that has a dependency on more than one "old" physical action. If an architect traces over an element and at the same time pays attention to another element near the first element, and notices the proximity between the two elements for the first time, this is an instance of unexpected discovery of the second type.

TABLE 1. Three distinct types of unexpected discoveries

type	definition		description
	behavior	dependent on	
type 1	"new" attention to a shape, size or texture	a single "old" physical action	discovery of a visual feature of an element
type 2	"new" attention to a relation	more than one "old" physical actions	discovery of a spatial or organisational relation among elements
type 3	"new" attention to an empty space among elements	implicit	discovery of an implicit space that exists in between elements

TABLE 2. Types of "old" physical actions

name	description of action
touching/tracing	touch or trace over a previously drawn element on a sketch
copying	copy a previously drawn element on a new sheet of paper from the sheet underneath
inspecting	pay attention to the existence of a previously drawn element

The third is the discovery of a space that exists in between previously drawn elements. This is so-called perception of figure-ground reversal, one of the characteristics of human perception. A famous example from psychology is perception of a single vase versus the contours of two human faces facing each other. Arnheim, under a chapter titled "solids and hollows" in his book (1977), has given many examples of this in the context of architectural design. The definition of this type is an exception; its dependency on "old" physical actions is implicit. Of course, when an architect mentions the discovery of a space of this sort, he or she must be paying attention to the existence of surrounding elements. But, it is not clear about which of the surrounding elements he or she is actually paying attention to. This implicitness did not hamper the codings of this type, because, while reporting, our architect pointed to the area of an implicit space on the TV screen where his sketching activities were replayed.

3.3 GOALS TO INVENT IMPORTANT ISSUES AND REQUIREMENTS

3.3.1 Distinction between Functional Actions and the Set-up of Goals

As a preparation for explaining how to code goals for S-invention, we will first describe the distinction between functional actions and conceptual actions in our coding scheme. **Functional** actions refer to actions of associating particular visuo-spatial features in sketches with meanings, functions or abstract concepts. For example, if an architect attends to a proximity between a plaza in front of the museum building and a street outside the property line, and thinks of a view from and to each other, the thought about the "view" is an instance of functional action.

Here, a question arises. Whenever an architect conceives of a "new" meaning, function or abstract concept associated with a particular visuo-spatial feature, should we code the instance as the S-invention of an important issue and requirement? The answer is no. What we mean by "the S-invention of important issues and requirements" has the following connotation; that is, an issue should be abstracted out of specific situations in sketches and become general enough to be carried through the entire design process as one of the primary design requirements. We assume that this act appears as the set-up of goals in design processes. In the example of "view" presented before, if the architect detaches the "view" from the particular proximity between the two regions, and says "let's always introduce a nice view from the street to the main feature of the museum", then we code this instance as the S-invention of an important issue and requirement. The architect has set up a goal to introduce a function, i.e. nice view from the street to the main feature of the museum, and may carry this design requirement through the subsequent process as one of his primary concerns.

The set-up of goals is one category of **conceptual** actions in our coding scheme. The other categories of conceptual actions are preferential or aesthetic evaluations and retrieval of knowledge or past similar cases (Suwa, Purcell and Gero, 1998). We will not go into the details of these other categories, because they are not relevant to the purpose of this paper.

3.3.2 Distinct Types of Goals and Their Examples

From intensive observation of the verbal protocol of our architect, we have classified the set-up of goals into distinct categories. Table 3 shows all the categories. First, there are four major categories; one is goals to introduce new functions at the current segment (Type 1 goals). The second is goals to resolve problematic conflicts that are detected in the current design (Type 2 goals). The third is goals to apply previously-introduced functions or arrangements in the current context (Type 3 goals). The fourth is repeated goals from a previous segment (Type 4 goals).

The first category is, in turn, divided into several subclasses. The first is goals to create functions that are listed in the initial requirements given to the architect at the beginning of the design session (Type 1.1). For example, in the design task which Suwa and Tversky (1997) gave to architects, they were asked to arrange not only museum building(s) but also green area(s) with pond(s), sculpture gardens, and a parking lot. So, whenever there is an instance that the architect intentionally introduced one of these functions at a new spot in his or her sketch in a way which had never been done before, we code it as Type 1.1 goal.

TABLE 3. Types of goals to invent new functions and issues

Type 1: goals to introduce new functions	
Type 1.1:	based on the given list of initial requirements
Type 1.2:	directed by the use of explicit knowledge or past cases
Type 1.3:	extended from a previous goal (subtypes: concretizing & broadening)
Type 1.4:	in a way that is not supported by knowledge, given requirements, or a previous goal
Type 2: goals to resolve problematic conflicts	
Type 3: goals to apply previously introduced functions or arrangements in the current context	
Type 4: repeated goals from a previous segment	

The second is goals triggered by retrieval of knowledge or past similar cases (Type 1.2). For example, if an architect mentions a piece of knowledge that a museum typically has restrooms near its entrance hall, and sets up a goal to create a "restroom" function, this is an instance of Type 1.2 goal.

The third is goals derived from a previous goal in an extended way (Type 1.3). One way of extension is to extend an issue dealt with in a previous goal into a more concrete specification. For example, at a fairly early phase of his design process, our architect set up a goal that, in order to guide people in a nice way from the parking lot to the main building of the museum, he should somehow control the flow of people after they park their car. Based on this goal, after a while, he set up another goal that he should arrange series of amusing objects, such as sculptures, water streams, or whatever, along the procession space from the parking lot to the building. The second goal is an extension of the first one; he specified in a concrete way how to "control the flow of people".

Another form of extension is in an opposite direction, i.e. to generalise the issue dealt with in a previous goal and bring it into a broader context. For example, after he had tentatively determined the position of an entryway from the outside street to the site of the museum, our architect set up a goal to decide the shape of the entryway. Then, suddenly, he began to question, from the concern of traffic interruption, whether or not it is really possible for visitors to enter the site from that spot, and set up another goal to consider the position of the entryway in the context of outside traffic. The concern of the second goal is broader than the first one, although still dealing with the issue of designing the entryway. This corresponds to what Lawson (1991) called "defining problems by escalation".

The fourth is goals to create new functions in a way that is not supported by knowledge, previous cases or experiences, initially given requirements, or previous goals (Type 1.4). We code goals to introduce new functions into this category, when we cannot find in the verbal protocol any supporting evidence of how the architect conceived of the functions. For example, our architect set up a goal that when visitors walk from the parking lot to the main building of the museum, they should be soaked in bright sunlight. It was for the first time in his design process that he talked about the relationship between the pathway and sunlight. Neither explicit knowledge nor past experiences were reported. Goals in this category must have derived from some justified reasons. In the above example, the architect sought justification in a reason such as "people don't want to walk in a dark place". We posit that reasons of this sort are fundamentally different in nature from explicitly articulated knowledge; these are not retrieved from an organized set of domain knowledge or past cases in memory, but rather constructed on the fly in a situated way. For this reason, we distinguish goals in this category from Type 1.2 goals.

An example of Type 2 goals is as follows. Our architect once decided to bring a water stream from the open plaza in front of the museum building into the entrance hall, as a means to guide visitors into the building in a cheerful way. Then, after a while, he noticed that water in the building may cause problems because humidity affects the artworks. But, because he thought that the idea of bringing water inside is still promising to produce a lively atmosphere, he set up a goal to search for a method to let artworks and water co-exist. In most cases, judgement of this sort that some aspects in the current design are problematic is mediated by domain knowledge. But, goals in this category are distinct from Type 1.2 goals; Type 1.2 goals are set up to introduce new functions as the corresponding knowledge or past cases prescribe, while knowledge just mediates the judgement done before the set-up of Type 2 goals.

An example of Type 3 goals is as follows. In the middle of his design process, our architect discovered that there is nothing physical on a line extending from an entryway into the site to the sculpture garden in the relatively far end of the site, and that a nice view to the cheerful sculptures could be the first experience for visitors when they drive through the entryway. The thought on this particular "view" per se is a functional action in our coding scheme, because it is bound by the particular spatial relation between the entryway and the sculpture garden, and thus still not abstracted out as a goal. Then, while he was sketching in a later sheet of paper, he remembered this, abstracted it into a general issue such as "create a nice view from entryway to a sculpture garden", and applied it to the situation at that moment.

3.3.3 Goals for the S-invention of Important Issues and Requirements

Which types of goals should be interpreted as instances of "the S-invention of important issues and requirements"? Obviously, Type 1.1 goals should not be included, because the functions introduced by them are given in the list of initial requirements, not invented during the process. Because we are interested in the exact moment when inventive issues or design requirements get born, we do not code Type 4 goals as instances of S-invention, even though it was an instance of invention when it was originally set up.

We assume that Type 3 goals should be omitted, too. Although it is for the first time during the design process that an architect sets up a goal to deal with the issue or requirement, the seed for this goal existed at a previous segment in the form of a particular function associated with particular visuo-spatial features. Therefore, we cannot say in a true sense that the general issue or requirement has been invented exactly at this moment. For the same reason we reject Type 4 goals, Type 3 goals should not be included.

The remaining goals, Type 1.2, Type 1.3, Type 1.4, Type 2, all satisfy the requirement for "S-invention"; the issue or requirement was born at that moment for the first time, although the ways it has occurred vary from time to time. Thus, we regard goals belonging to these four categories as instances of "the S-invention of important issues and requirements".

4. Results

4.1 CODING OF UNEXPECTED DISCOVERIES AND GOALS

The entire protocol of our architect contained 340 segments. For each segment, we identified and coded instances of unexpected discoveries and goals. The entire protocol contained 608 perceptual actions, out of which 173 were unexpected discoveries. The fact that a significant portion of the perceptual actions, 28.5%, belonged to unexpected discoveries clearly shows its importance in the design process of our architect. Out of the 173 instances of unexpected discoveries, 38 belonged to the discovery of visual features of elements, 106 to the discovery of spatial or organisational relations among elements, and 29 to the discovery of implicit spaces.

The entire protocol contained 240 instances of the set-up of goals. Table 4 shows the number of occurrences of each type of goal. The most frequent goals were Type 1 goals, the ones to introduce new functions. 60% of the total belonged to this type. This is not surprising because the task of a designer is to introduce and arrange many new functions so that they make a coherent harmony under given requirements. The second most frequent goal was Type 3, the one to apply previously-introduced functions and arrangements in the current context. The design of our architect developed as he went through many sheets of sketches. He produced 13 pages of sketches. This suggests that one important aspect of his design was to remember functions and/or arrangements which he had adopted previously, abstract them out of the particular contexts of previous segments and apply them in the current context.

TABLE 4. The numbers of occurrences of distinct types of goals

Types	Number of goals identified	Percentage to the total (%)
Type 1 total	144	60.0
Type 1.1	32	13.3
Type 1.2	40	16.7
Type 1.3	43	17.9
Type 1.4	29	12.1
Type 2	16	6.7
Type 3	54	22.5
Type 4	26	10.8
G_{inv} (Type 1.2 + Type 1.3 + Type 1.4 + Type 2)	128	53.3
total	240	

We had 128 instances of goals for S-invention, i.e. the sum total of Type 1.2, Type 1.3, Type 1.4, and Type 2. This corresponds to 53.3% of the total occurrences of goals. The fact that about a half of the goals relate to the act of inventing important issues or requirements suggests its importance in the design process of this architect. Further, goals related to S-invention were much more frequent than goals to introduce new functions based on the given list of requirements, 144 vs. 32. This clearly shows one characteristic of conceptual design processes using freehand sketches. Designers must satisfy a given list of requirements in their designs. But, more than that, they should be able to invent design issues or requirements through interaction with sketches, sometimes by using explicitly articulatable knowledge, and sometimes by constructing justifiable reasons on the fly.

4.2 CORRELATION BETWEEN UNEXPECTED DISCOVERIES AND GOALS FOR INVENTIONS

Is causality from unexpected discoveries to the S-invention of design issues or requirements a phenomenon throughout the entire design process? It may be so, or may be only during certain periods. How should we identify the periods in which there is causality? We conjectured as follows. If there is any significant causality in certain periods, then the frequencies of the occurrences of both might increase or decrease in a similar way in those periods. Thus, we have examined whether or not the frequencies of the occurrences of both changed over time in a similar way with each other. In the following sections, for simplicity, we will denote unexpected discoveries as UXD, and goals for S-invention as G_{inv}.

For this examination, using the segment-by-segment changes of frequencies of UXD and G_{inv} would not be a good idea, because our concern is to roughly examine the tendencies in which the frequencies of both change over time. For each page of the sketch that the architect produced, we chunked every five segments from the beginning, and thereby calculated the sum total of the occurrences of UXD and G_{inv} in each 5-segment block. Since the number of segments in each page is not necessarily a multiple of five, the last block in each page could consist of 1, 2, 3, 4 or 5 segments. If the number of segments in the last block is too small, the situation would be close to an examination of segment-by-segment changes, and thus not desirable. Therefore, we did the following modification. If the last block of each page turned out to consist of 1 or 2 segments, we merged these segments into its immediately previous block, so that the resulting last block became to consist of 6 or 7 segments. If the last block of each page turned out to consist of 3, 4 or 5 segments, we adopted it as the last block as it was. So, although most blocks throughout the entire protocol consist of 5 segments, only the last block of each page consist of 3, 4, 5, 6 or 7 segments. Why should we stick to this modification instead of just simply dividing the entire protocol into 5-segment blocks? It is because we assume that behaviors of a designer in the last portion of each page and those in the beginning part of the next page are different in nature, and thus should not be merged into each other.

In Section 3.3.3, we defined G_{inv} as goals belonging to Type 1.2, Type 1.3, Type 1.4, or Type2. Although all relate to the S-invention of design issues or requirements, however, we assume that Type 1.2 goals and the remaining three types of goals are slightly different in nature. Type 1.2 goals are set up according to the prescription of what a piece of explicitly articulated knowledge or past cases dictates, while the other types are not. Therefore we determined to investigate the following two correlations. One is correlations between UXD and goals belonging to any of the four types (we will denote this as G_{inv1}). The other is correlations between UXD and goals belonging to either Type 1.3, Type 1.4, or Type 2 (we will denote this as G_{inv2}).

Figures 1 and 2, respectively, show the block-to-block changes of the frequencies of occurrences of UXD and G_{inv1}, and of UXD and G_{inv2}, throughout the entire protocol of our architect. The horizontal axis is the block number. We had 68 blocks for the entire 340 segments. The vertical axis is the frequency normalized as follows. The sum total of occurrences in each block is divided by the number of segments in the block, to get the frequency per segment in the block, *F*. Then, we calculate the average, F_{avg}, and standard deviation, F_{std}, of *F* over the entire 68 blocks. As the vertical value, we use $(F-F_{avg})/F_{std}$. We did this because, in examining correlations between UXD and G_{inv1} or G_{inv2}, it might be better to compare by removing the magnitudes of frequency that are specific to both.

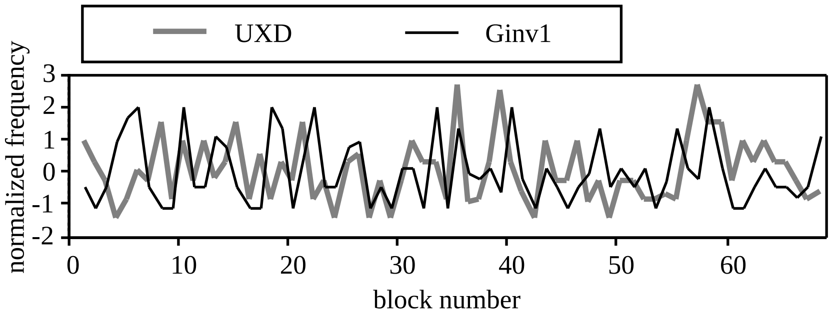

Figure 1. Block-to-block changes of frequencies of both UXD and G_{inv1}

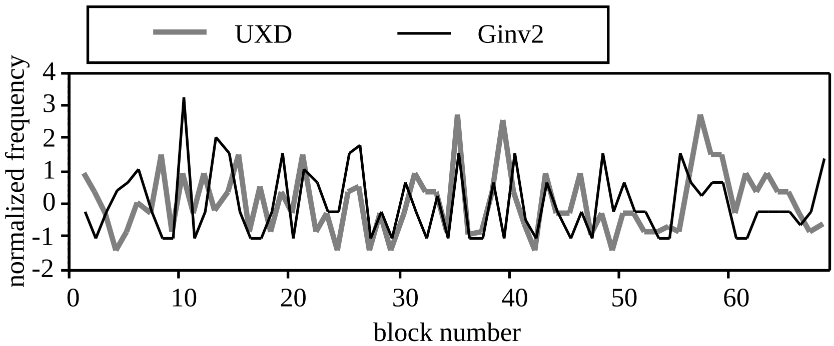

Figure 2. Block-to-block changes of frequencies of both UXD and G_{inv2}

Based on the data on the normalized frequencies of UXD and G_{inv1} (or G_{inv2}) in each block, we examined whether or not there are any portions of the entire design process in which the frequencies of both changed in correlation with each other. We carried out statistical analyses in the following manner. First, for each block, we calculated the difference of the normalized frequency from its immediate previous block. Then, we identified a period of consecutive blocks, under the constraint that the length of the period should be at least three blocks, in which the sequence of block-to-block differences of UXD correlate with that of G_{inv1} (or G_{inv2}), i.e. UXD and G_{inv1} (or G_{inv2}) increased or decreased in a similar manner. We did this analysis by conducting χ-square test on the pair of sequences of differences of UXD and G_{inv1} (or G_{inv2}). We picked up only periods of consecutive blocks for which correlation is statistically valid with a certainty of more than 90%. The reason why we picked up a period of blocks whose length is at least three is that, by doing so, we can eliminate the cases in which UXD and G_{inv1} (or G_{inv2}) happened to increase or decrease to a similar extent for a single transition from one block to the next block. More than one period of consecutive blocks that satisfy these conditions are expected to be identified in the entire design process, and each period of blocks is regarded as a period of correlation between UXD and G_{inv1} (or G_{inv2}).

Figure 3 shows, for the pairs of UXD and G_{inv1} and of UXD and G_{inv2}, the periods of correlation. The horizontal axis is the block number, representing the time frame of our architect's process. The horizontal bars represent periods of correlation. The number written near each bar is the identification number of the period, corresponding to each period number in Table 5. Table 5 shows the statistical data for each period of correlation: the duration of the correlation in terms of the number of consecutive blocks, χ-square value, and a certainty.

UXD and G_{inv1} correlated with each other in 26 blocks in total, which corresponds to 127 segments. This covers 37% of the entire process. The periods of correlation existed from the middle of the entire process to the end. On the other hand, UXD and G_{inv2} correlated with each other in 36 blocks in total, which corresponds to 177 segments. This covers 52% of the entire process. The periods of correlation cover almost those periods for the former pair, and additionally some other periods in which the former pair does not correlate.

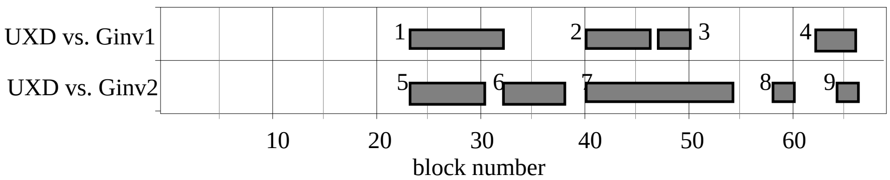

Figure 3. The periods of correlation between UXD and G_{inv1}, and between UXD and G_{inv2}.

TABLE 5. The statistical data of periods of correlation

period number	duration of correlation	χ-square value	certainty
1	10	3.07	p>0.9
2	7	1.85	p>0.9
3	4	0.58	p>0.9
4	5	0.09	p>0.995
5	8	1.56	p>0.975
6	7	1.89	p>0.9
7	15	5.54	p>0.975
8	3	0.001	p>0.995
9	3	0.027	p>0.975

We have obtained the following insights. First, although correlations between UXD and G_{inv1} (or G_{inv2}) do not necessarily occur in the entire process, correlations between UXD and G_{inv} are actually a major phenomenon; they cover a significantly large part of the entire design process. This suggests the possibility of causality between unexpected discoveries and goals for S-invention in significantly large parts of the design process. Second, the periods of correlation between UXD and G_{inv2} almost cover, and are significantly longer than, the periods of correlation between UXD and G_{inv1}. This suggests that unexpected discoveries may have a stronger causal relation to Type 1.3, Type 1.4, or Type 2 goals than to Type 1.2 goals.

4.3 CAUSALITY BETWEEN UNEXPECTED DISCOVERIES AND GOALS FOR INVENTIONS

The analysis in the preceding section was on the basis of blocks; we identified the periods of correlation between the occurrences of unexpected discoveries and those of goals for S-invention, thinking that correlations between both might imply causal relations between both. As the next pursuit, we examined, more precisely using the granularity of segments, whether there are really causal relations between UXD and G_{inv1} and between UXD and G_{inv2}, and if any, what kinds of causal relations. We did this, for each pair, only for the periods of blocks in which there was a correlation between the pair, because no causal relations between the two are expected for the periods in which there was no correlation. Hence, 127 segments for the pair of UXD and G_{inv1}, and 177 segments for the pair of UXD and G_{inv2}, were the targets of this examination.

How did we examine this? We will explain it, taking as an example the examination of whether there is a causality from UXD to G_{inv1}. We classify the entire target segments, i.e. 127 segments, into the following four classes in terms of the occurrences of UXD;

(1) class 1: segment which has no UXD but whose immediately previous segment had at least one UXD (we call this "next of but not together with UXD"),

(2) class 2: segment whose immediately previous segment had no UXD but which has at least one UXD ("together with but not next of UXD"),

(3) class 3: segment both of which and whose immediately previous segment has at least one UXD ("next of and together with UXD"),

(4) class 4: the remaining segments.

If there is no explicit causality from UXD to G_{inv1}, then the occurrences of instances of G_{inv1} should be distributed to each of the four classes of segments in proportion to the ratio of the number of each class to the total. On the contrary, if there is an explicit causality, more instances of G_{inv1} should be distributed partially to particular kinds of segments.

Table 6 shows the number of segments belonging to each class and the occurrences of G_{inv1} at each class of segments. The difference from the expected occurrences at each class of segments, which is calculated from the ratio of the number of segments belonging to each class to the total number of segments, is also shown.

TABLE 6. The occurrences of G_{inv1} in the four classes of segments categorised in relation to the occurrences of UXD, during periods of correlation between G_{inv1} and UXD

classes of segments	number of segments	G_{inv1}	
		number of occurrences	difference from the expected value
class 1: next of but not together with UXD	29	12	+3.4
class 2: together with but not next of UXD	28	8	- 0.4
class 3: next of & together with UXD	17	5	- 0.1
class 4: other	53	13	- 2.9
total	127	38	
partiality of distribution			no ($\chi^2(3) = 1.80$, $p>0.5$)

According to the statistics, instances of G_{inv1} occurred more frequently at class 1 segments than an expected level. However, the distribution of instances of G_{inv1} over the four classes of segments is not partial in a statistically significant manner, $\chi^2(3) = 1.80$ ($p>0.5$). Then, we combined segments belonging to class 1 and class 3 into a single class, and segments belonging to class 2 and class 4 into another. The former is the class of segments whose immediately previous segment has at least one UXD, while the latter is the class of segments whose immediately previous segments does not have a UXD. The distribution of instances of G_{inv1} over the two new classes, however, is not partial statistically, $\chi^2(1)=1.19$ ($p>0.25$); there is no such tendency that, when a UXD occurred at a segment, G_{inv1} is more likely to occur at the next segment. In any of the several ways of combination, we did not find a partial distribution of G_{inv1}. These results suggest that there is no explicit causality from UXD to G_{inv1}.

Table 7 shows the statistics about the causality from UXD to G_{inv2}. Instances of G_{inv2} occurred more frequently at class 1 segments than an expected level, and slightly more at class 3 segments. The distribution of instances of G_{inv2} over the four classes of segments, however, is not partial in a statistically significant manner, $\chi^2(3) = 3.60$ ($p>0.25$).

TABLE 7. The occurrences of G_{inv2} in the four classes of segments categorised in relation to the occurrences of UXD, during periods of correlation between G_{inv2} and UXD

classes of segments	number of segments	G_{inv2} number of occurrences	G_{inv2} difference from the expected value
class 1: next of but not together with UXD	38	13	+4.6
class 2: together with but not next of UXD	38	7	- 1.6
class 3: next of & together with UXD	22	6	+1.0
class 4: other	79	14	- 3.9
total	177	40	
partiality of distribution			no ($\chi^2(3) = 3.60$, $p>0.25$)

However, when we compare segments belonging to class 1 and class 3 with those belonging to class 2 and class 4, a different pattern appeared. Table 8 shows that instances of G_{inv2} were, in a statistically significant manner, more likely to occur in segments belonging to class 1 or class 3 that in those belonging to class 2 or class 4, although the tendency is not extremely strong, $\chi^2(1) = 3.30$ (p=0.08). In other words, there is a causal relation that, when an unexpected discovery has occurred at a segment, then an instance(s) of G_{inv2} is likely to occur at the next segment. Remember that this is the result of close examination of the target 177 segments, 52% of the entire process. This indicates that, for this architect, unexpected discoveries was one of the driving forces for the occurrences of G_{inv2}.

TABLE 8. The occurrences of G_{inv2} in the combined classes of segments categorised in relation to the occurrences of UXD, during periods of correlation between Ginv2 and UXD

classes of segments	number of segments	G_{inv2} number of occurrences	G_{inv2} difference from the expected value
class1 + class3: next of UXD	60	19	+5.4
class2 + class4: other	117	21	- 5.4
total	177	40	
partiality of distribution			yes ($\chi^2(1)$ = 3.30, p=0.08)

Further, we have examined whether there was causality from G_{inv1} to UXD and from G_{inv2} to UXD, too. Tables 9 and 10 show the statistics for each pair, respectively. The target segments, i.e. 127 segments for the former pair and 177 for the latter, were classified into the four classes in terms of occurrences of G_{inv1} (or G_{inv2}). Table 9 shows that UXDs occurred more frequently at class 1 segments, such segments whose immediately previous segment had at least one G_{inv1} but which do not have, than an expected level. Statistically, however, the distribution of instances of UXDs over the four classes of segments is not partial, $\chi^2(3)=3.76$ (p>0.25). Further, no matter how we combined classes of segments, we did not find in a statistically significant way a partial distribution of UXDs. These results suggest that there is no explicit causality from G_{inv1} to UXD.

On the other hand, we found a strong causality from G_{inv2} to UXD. Table 10 shows that UXDs occurred more frequently at class 1 segments and less at class 4 segments. The distribution of UXDs over the four classes of segments were partial in a statistically significant way, $\chi^2(3)=8.31$ ($p<0.05$). This effect appeared more salient, when we compare the combined classes of class 1, 2 and 3 with class 4. Table 11 shows a statistically strong causal relation that UXDs were likely to occur more frequently at such segments either of which or whose immediately previous segment had at least one G_{inv2}, $\chi^2(1)=6.59$ ($p=0.01$).

Table 9: The occurrences of UXD in the four classes of segments categorised in relation to the occurrences of G_{inv1}, during periods of correlation between G_{inv1} and UXD

classes of segments	number of segments	UXD: number of occurrences	UXD: difference from the expected value
class 1: next of but not together with G_{inv1}	26	17	+5.7
class 2: together with but not next of G_{inv1}	26	9	- 2.3
class 3: next of & together with G_{inv1}	8	3	- 0.4
class 4: other	67	26	- 3.0
total	127	55	
partiality of distribution			no ($\chi^2(3) = 3.76$, $p>0.25$)

Table 10: The occurrences of UXD in the four classes of segments categorised in relation to the occurrences of G_{inv2}, during periods of correlation between G_{inv2} and UXD

classes of segments	number of segments	UXD: number of occurrences	UXD: difference from the expected value
class 1: next of but not together with G_{inv2}	31	22	+8.5
class 2: together with but not next of G_{inv2}	30	15	+1.9
class 3: next of & together with G_{inv2}	8	4	+0.5

class 4: other	108	36	- 11.0
total	177	77	
partiality of distribution			yes ($\chi^2(3) = 8.31$, $p<0.05$)

Table 11: The occurrences of UXD in the combined classes of segments categorised in relation to the occurrences of G_{inv2}, during periods of correlation between G_{inv2} and UXD

classes of segments	number of segments	G_{inv2}	
		number of occurrences	difference from the expected value
class1 + class2+ class3: next of and together with G_{inv2}	69	41	+11.0
class4: other	108	36	- 11.0
total	177	77	
partiality of distribution			yes ($\chi^2(1) = 6.59$, $p=0.01$)

5. Discussions and Implications

In about a half of the entire design process of our architect, i.e. 52%, we found causal relations from unexpected discoveries to goals for S-invention except knowledge-derived ones, and from the latter to the former. When at least one unexpected discovery has occurred at a segment, an instance of those goals is likely to occur at the next segment, although this tendency was not extremely strong. Further, when at least one instance of those goals has occurred at a segment, an unexpected discovery is likely to occur at the same or the next segment. This was a strong tendency. We interpret this as follows. There were strong feedback loops between unexpected discoveries and the S-invention of design issues or requirements. When this architect discovered a visuo-spatial feature in his sketch in an unexpected way, it encouraged him to invent important issues or requirements that are relevant to the visuo-spatial feature. Further, the S-invention of important issues or requirements made himself committed more deeply to the area of a sketch relevant to the invented issues or requirements, and thus encouraged other unexpected discoveries in those areas.

As opposed to our original assumption that the S-invention of design issues or requirements should consist of knowledge-derived goals, goals extended from a previous one, unsupported goals and goals to resolve problematic conflicts, the inclusion of knowledge-derived goals obscured causal relations to and from unexpected discoveries. This means that while the last three types of goals occur in strong correlation with unexpected discoveries, knowledge-derived goals do not. We interpret this as follows. There are at least two distinct ways in which design issues or requirements are invented during the process. One way is to retrieve explicit knowledge or past cases and produce issues or requirements as the knowledge or the cases prescribe. In this case, the retrieval of the knowledge of past cases itself is the core of invention. The other way is to conceive of issues or requirements, not by explicit use of knowledge or past cases but by some sorts of justifications or reasons which are spontaneously constructed at the moment. In this case, unlike the first way of invention, it is an open question how these justifications or reasons occur in a designer's mind. It appears that they are not retrieved from an explicitly articulatable set of knowledge or past cases but must be constructed on the fly during the process. And unexpected discoveries seem to be involved in this process.

This finding is strongly supportive evidence of recently prevailing view that designers construct design ideas on the fly in a situated way, responding to visuo-spatial features of the physical setting in which they sketch (Suwa, Gero and Purcell, 1998; Gero, 1998). Design ideas, e.g. S-invention in this paper, occur at a moment during the design process, as a result of what has occurred so far in the design setting, i.e. what the designer sketched, perceived and thought of. The same ideas would not necessarily come to appear even if the same designer works on the same task in future, because the appearances of his or her sketches in the second trial and what he or she perceives in them would be different.

Weisberg (1993) argued that the necessary condition for creative solutions in a problem-solving process is discontinuous jumps during the process. He raised analogical reasoning as one of the promising means to introduce discontinuous jumps in the process. The findings in this paper indicate that the S-invention of important issues or requirements epitomises one way to introduce discontinuity in the context of design.

What implications do these findings have? There has been a belief that the invention of design issues or requirements during the process is the key to gaining creative outcomes in the end. Supporting evidence for this comes from the experiments in which "problem-finding" behaviors of art students in a drawing task were studied (Getzels and Csikszentmihalyi, 1976). They found that the more frequently students displayed "problem-finding" behaviors during the task, the more creative their final drawings were rated.. Further, there was a strong correlation between their "problem-finding" tendency in the experimental task and the degree of success in later years as professional artists. Although pioneering in clarifying the correlation between "problem-finding" behaviors and creativity, this study has left the following question in the pedagogical sense unanswered. How do students, whether in art or in design, acquire the ability to behave in a "problem-finding" manner, i.e., in our terminology, to invent issues and requirements during the process? How should educators encourage students to do so? The findings in this paper have given a basis for the answer. That is, the following advice may work well for a student. "Let's try to discover implicit visuo-spatial features in your sketches that you did not intend when you drew, and think of design issues or requirements associated with those features". Because it clearly sets the goal for students, i.e. intentional search for unintended features in their sketches, this sort of instruction is more easily carried out and thus more practical than an abstract advice such as "just behave in a problem-finding manner".

6 Future Work

First, further analyses are needed covering more practising architects, to see whether the findings in this paper are a general tendency common with practising architects. Further, comparisons between practising architects and students in the frequencies at which unexpected discoveries and goals for inventions occur and the ways in which both cause each other will be productive to gain precise directions about how to educate students toward creative outcomes.

Second, examinations of whether there are causal relations between goals for inventions and particular classes of physical or perceptual actions are needed. For example, attending to a class of visual features such as shapes or a class of spatial relation such as proximity, whether newly or revisiting these visuo-spatial features attended before, might have a particular importance for designers. Particular classes of physical actions, such as tracing over previous depictions, might correlate with unexpected discoveries and/or goals for inventions. This examination is expected to provide more precise understanding of the ways in which designers' actions are situated in the physical setting they draw and perceive.

7. Conclusion

The aim of this research was to verify the hypothesis that unexpected discoveries may become the impetus for the invention of design issues or requirements, seeking empirical evidence through protocol analysis. First, we defined unexpected discoveries as a class of perceptual actions; actions of attending to visuo-spatial features of previously drawn elements. Then, we assumed that four types of design goals are the instances of the S-invention of issues or requirements; goals derived from explicit knowledge, goals extended from a previous goal, unsupported goals, and goals to resolve problematic conflicts.

We analysed the cognitive processes of a practising architect in a design session. We identified the portions of the entire design process in which the frequency of the occurrences of unexpected discoveries and that of goals for S-invention changed over time in correlation with each other. Then, for the periods of correlation identified, we more closely investigated what kinds of causal relations both have to each other. We have obtained the following findings. First, it was a composite of the three types of goals excluding ones derived by explicit knowledge that had causal relations with unexpected discoveries. This suggests that there are at least two distinct ways in which design issues or requirements are invented during the process; one is in a knowledge-directed way, and the other is in a more implicit way without mediation of explicit knowledge or past cases. Unexpected discoveries involve the latter means of invention. Second, the periods of the causal relations covered 52% of the entire process. This suggests that causality between the two was a major phenomenon in the process of this architect. Third, causal relations were bi-directional. When an unexpected discovery has occurred at a segment, a goal belonging to either of the three types is more likely to occur at the next segment. When a goal belonging to either of the three types has occurred at a segment, an unexpected discovery is more likely to occur at the same or the next segment. Not only unexpected discoveries became the driving force for S-invention, but also the occurrences of S-invention, in turn, made the architect more deeply committed to the areas of a sketch relevant to the invention and caused new unexpected discoveries in those areas. These form supporting evidence that design is a situated act; design ideas occur as a result of what has occurred so far in the design setting in which the designer is.

These findings significantly imply that encouraging design students to have the disposition of intentional search for unintended features in their own sketches might be a good basis for education toward creative outcomes.

Acknowledgment

This research has been funded by the Australian Research Council, Grant Number A89601894.

References

Arnheim, R.: 1977, *The Dynamics of Architectural Form*, University of California Press, Berkeley.

Boden, A.: 1990, *The Creative Mind: Myths and Mechanism*, Weidenfeld and Nicolson, London.

Gero, J. S.: 1998 Conceptual designing as a sequence of situated acts, *in* I. Smith (ed.), *Artificial Intelligence in Structural Engineering,* Springer, Berlin, pp. 165-177.

Gero, J. and McNeill T.: 1998, An approach to the analysis of design protocols, *Design Studies*, **19**(1): 21-61.

Getzels, J. W. and Csikszentmihalyi, M.: 1976, *The Creative Vision: A Longitudinal Study of Problem Finding in Art*, John Wiley & Sons, New York.

Goel, V.: 1995, *Sketches of Thought*. MIT Press, Cambridge.

Goldschmidt, G.: 1991, The dialectics of sketching, *Creativity Research Journal,* **4**(2): 123-143.

Goldschmidt, G.: 1994, On visual design thinking: the vis kids of architecture, *Design Studies*, **15**(2): 158-174.

Lawson, B.: 1990, *How Designers Think*, Butterworth Architecture, Oxford.

Robbins, E.: 1994, *Why Architects Draw*, MIT Press, Cambridge.

Schon, D. A. and Wiggins, G.: 1992, Kinds of seeing and their functions in designing, *Design Studies*, **13**(2): 135-156.

Stenning, K. and Oberlander, J.: 1995, A cognitive theory of graphical and linguistic reasoning: logic and implementation, *Cognitive Science*, **19**(1): 97-140.

Suwa, M., Gero, J. and Purcell, T.: 1998, The roles of sketches in early conceptual design processes, *Proceedings of Twentieth Annual Meeting of the Cognitive Science Society*, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp. 1043-1048.

Suwa, M., Purcell, T. and Gero, J.: 1998, Macroscopic analyses of design processes based on a scheme for coding designers' cognitive actions, *Design Studies*, **19(**4): 455-483.

Suwa, M. and Tversky, B.: 1997, What do architects and students perceive in their design sketches?: A protocol analysis, *Design Studies* **18**(4): 385-403.

van Someren, M. W., Barnard, Y. F. and Sandberth, J. A. C.: 1994, *The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes*, Academic Press, London.

Weisberg, R. W.: 1993, *Creativity: Beyond the Myth of Genius,* W. H. Freemand, New York.

This paper is a copy of: Suwa, M., Gero, J. S. and Purcell, T. (1999) Unexpected discoveries and s-inventions of design requirements: A key to creative designs, *in* J. S. Gero and M. L. Maher (eds), *Computational Models of Creative Design IV*, Key Centre of Design Computing and Cognition, University of Sydney, Sydney, Australia, pp. 297-320