The form of design processes: a protocol analysis study

Michael Eckersley

University of Maryland, Department of Design, 1401 Marie Mount Hall, College Park, Maryland 20742, USA

Protocol analysis (PA) methodology is discussed in light of its potential for empirical analysis of design problem-solving behaviour. A literature review highlights previous protocol analysis studies in design, with particular attention to Akin's[1] PA case study. Presented is a new model-based PA method whereby the verbal behaviour of designers is analysed, quantified, and statistically manipulated to reveal a unique and scientifically rigorous facsimile of problem-solving processes. A recent study based on the above method is outlined, and results of the study are reported.

Keywords: protocol analysis, cognition, verbal behaviour

The efficient and effective ordering of entities in space and/or time is a particularly valued capability in modern society. The result of such an ordering process is that of a designed artefact which serves to meet a perceived need or solve a given problem[2]. The more difficult and ill-structured the problem, usually the more complex is the path constructed to solve it. The general architecture of that path is the focus of the present study.

Because it is impossible to see into the black box of the designer's mind, it has been difficult for those studying cognitive processes to ascertain how one actually processes information and exploits knowledge to formulate the designed artefact. The goal of this research was to utilize the psychological method of protocol analysis (thereby adopting its basis of assumptions about human cognition) to analyse the verbal behaviour of a small group of designers engaged in problem-solving.

On one level the scale of the study was small, in that a non-random sample of five designers was utilized, and that no experimental treatment was given nor controlled for. However, the significance of the study lies in its being the first of its kind to

- chart design problem-solving behaviour over time across eight variables
- utilize specially developed computer programs (CODEPRO & CODESYN) to streamline the encoding and analysis of verbal protocols
- use a sample of three encoders to construct a more veridical interpretation of protocols
- allow for a modest testing of experimental hypotheses

The hypothesis tested in three ways was simply that designers vary significantly in the nature and amount of information processed during problem-solving. A secondary informal hypothesis was that stages in problem solving would be evident in the analysis of protocol data.

*The author would like to acknowledge the assistance of the IBM-Fulcrum CADD Lab in the Department of Design at The University of Maryland, and to thank those subjects, encoders, and programmers who assisted in the study.

REVIEW OF RELATED LITERATURE

Problems and problem-solving

Thomas, Lyon, and Miller[3] define problem-solving to be 'the activity by which a person (or other system) tries to achieve a goal when the person (or other system) has no existing (practical) algorithmic procedure for reaching the goal' (p. 3). Gagne's[4] description of problem-solving still contains the essence of contemporary thought on the subject. He suggested that problem-solving begins with a stimulus situation consisting of certain explicit or implicit instructions which establish the sets and define the goal. With that follow a series of phases:

- reception of the stimulus situation
- concept invention
- determination of the course, or courses of action
- decision-making when two or more courses are available, each appearing to provide adequate outcomes
- verification, where information of outcomes is fed back to the problem-solver for verification

Reitman[5] differentiated among problems in terms of their capacity for definition. He termed 'well-defined' those problems exhibiting well specified initial conditions, necessary operations, and goals. 'Ill-defined' problems on the other hand, according to Reitman, exhibit no such inherent specificity. Eastman[6] offered that many problems are ill-defined and therefore require subjective specification from the problem-solver, in terms of conditions, operations, and goals. Design problems are inherently ill-defined[7], and as such possess poorly specified initial conditions, allowable operations and goals. Simon[8] argues that such problems require an extremely large base of relevant knowledge to be effectively solved, and yet, when reduced to a series of sub-problems, are as solvable as well-structured problems.

Protocol analysis methodology

Protocol analysis[9,10] is a research methodology based on the psychological theory of information processing[6,11]. Intrinsic to information processing theory is the notion that thought is both the process and product of information processed by the brain. Since thought is not a directly observable activity, introspective and retrospective accounts of human thought have rightly been regarded as unreliable data for scientific enquiry. However, a small but emerging school of thought in psychology proposes that under controlled conditions, individuals (when trained to concurrently verbalize their thoughts) can reveal a remarkably accurate picture of their cognitive processes while engaged in problem-solving. The subsequent analysis of such problem-solving sessions (or protocols) is the substance of protocol analysis methodology. Simply stated, PA involves the recording of problem-solvers' concurrent verbalizations under controlled conditions, and the later assignment (or encoding) of portions of verbal protocol to previously defined categories in a model-based theory of cognition.

What protocol analysis offers the community of design scientists is a potentially effective method for the controlled observation and experimental analysis of design problem-solving behaviour. For a detailed review of PA theory, see Ericcson and Simon[9].

The essence of the theory rests on the following premises:

- Verbalize cognitions can be described as states that correspond to the contents of short-term memory (i.e. to the information that is in the focus of attention)
- Information vocalized is a verbal encoding of the information in short-term memory
- Verbalization processes are initiated as a thought is heeded
- Verbalization is a direct encoding of the heeded thought and reflects its structure
- Units of articulation can correspond to integrated cognitive structures
- Pauses and hesitations are good predictors of shifts in processing of cognitive structures

Implied in the theory is that information used to solve problems is of various kinds (e.g. inferential, hypothetical, recollected, evaluational, etc.). Further, it is assumed in the present study that such varietal information can be identified in verbal data, encoded as to type, and subjected to various statistical and graphing procedures.

Protocol analysis methodology in design research

Few studies to date have utilized PA to assess cognitive processes in design. Akin[1], conducting the first PA study recorded in the design literature, attempted to make explicit the intuitive problem-solving behaviour of a single architect engaged in a complex design problem. Akin's assumption was that his subject's self-reporting would lack completeness and veridicality. Therefore, he imposed a codification system based on work by Miller *et al.*[12] to lend meaning to the protocol data. Akin's codification system relied on the assumption that design consists essentially of a hierarchy of plan-directed behaviours, or 'schemata'. The schemata included instantiation, generalization, enquiry, inference, representation, goal-definition, specification and integration.

The significance of Akin's study is its attention to systematic analysis of the designer's overt physical and verbal behaviours. Certain weaknesses, however, are also to be found. And whereas some of the weaknesses were acknowledged by Akin himself, they are not trivial, and call into question many of his conclusions. A few limitations of the study were that:

- it was a case study of one individual, utilizing neither a means of control nor an experimental treatment
- the examiner actively interacted with the subject during the session, possibly biasing the results
- it failed to capture the temporal nature of problem-solving

Ullman, Stauffer and Deitterich[13] have more recently utilized PA to detail the behaviour of six individuals engaged in a real-world mechanical engineering problem. Their guiding motivations were to

- shed light on how mechanical engineers actually solve problems
- improve process efficiency and design product quality
- apply artificial intelligence (AI) methods to development of computer-aided design (CAD) tools
- develop more natural designer-CAD interfaces

Although their study was not experimental (in the truest sense), and approximately 80% of the potentially rich verbal data were thrown out early on, its methods were conceptually rigorous in highlighting essential events of the problem-solving process. They were also reasonably successful in modelling the temporal development of a real-world design solution. Ullman, Stauffer and Deitterich concluded that designers progress from systematic to opportunistic behaviour as the design evolves, and that sketches and drawings play a critical role in organization of the process itself.

Of note also is recent research which utilizes PA as a tool in the engineering of software[14,15]. The general trend in these reports (as well as that of Ullman *et al.*[13]) is to use PA to model expertise in a given domain for the subsequent design of knowledge-based (expert) systems.

DESIGN OF THE STUDY

Subjects

The sample ($N=5$) in the present study consisted of two experienced, practicing interior designers, and three junior-level interior design students at The University of Maryland's Department of Design. Interior designers are here distinguished from interior decorators. The former are accredited professionals trained in space-planning and in the design of architectural interiors; the latter are not. The Subjects were volunteers who donated their time for the purposes of this research. The study was conducted in March 1985.

Method

From careful observation of videotaped design protocols generated by individuals engaged in a particular space-planning problem, eight verbalization-types were identified (with varying degrees of frequency) and operationally defined. They include:

- Literal copy
  Exact or nearly exact verbal copy of a problem statement
- Paraphrased copy
  Verbalization which captures the basic content of a problem statement
- Inference
  Higher order conclusions, assertions, propositions, or justifications not given in the problem statement but generated by the problem-solver
- Intention/plan (future-related inference)
  Verbalization which indicates a decision to proceed upon an intended course of action dealing with the problem or part of the problem
- Move
  Statement implying the actual movement of characters
- Search
  Verbalization (sometimes in question form) indicating a need to gather information before acting on the problem, or a portion of the problem
- Specific assessment
  Assessment, comparison, or value judgement concerning the configuration of 1, 2, or 3 characters on the board
- General assessment
  Assessment, comparison, or value judgement concerning the general office layout
- None of the above
  Verbalization so unique as not to fit into any one of the above categories

A model (Figure 1) was developed to represent (globally) the design task-related verbalizations in a hierarchy of information structures. The model shows the eight types (i.e. LC, PN, IN, IP, MO, SE, SA, GA) as a subset of all verbalizable cognitions. Verbalizable cognitions are, in turn, shown as a subset of all cognitions passing through the short-term (STM) and working memories (WM) of the subject.

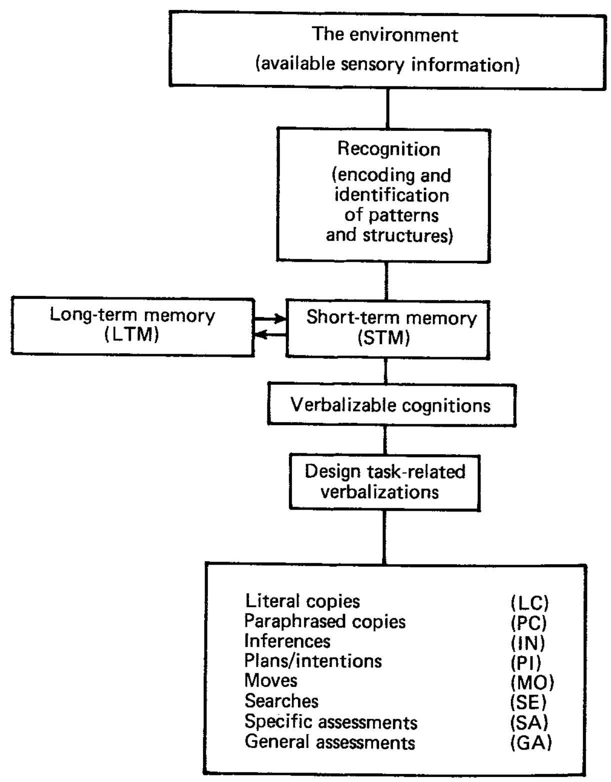


Figure 1. Model of information structures in design

The design problem

The design problem (Figure 2) itself was adapted from one conceived by Carroll, Thomas and Malhotra[16], wherein the space-designer must move seven characters around a schematic office space to accommodate 19 (often conflicting) functinal requirements. The problem consisted of three bipolar issues involving Office Function, Prestige, and Adjacency Preferences. The fourth central issue involved the Compactness/Cost of Offices. A cover story for the problem was also presented to the Subjects explaining the nature of the task they were to perform. The desire of the author was to control, in some measure, the structure of the problem so as to ensure a certain baseline of comparison between problem-solvers. By doing so, it was expected that the verbal behaviour generated would share some characteristics, including a somewhat common vocabulary of terms. Subsequent analysis of the protocols confirmed this expectation.

Data collection

Each subject received videotaped instructions describing the nature of the problem and their goal as designers to solve it. To train subjects to think-aloud, three brief practice problems (similar to those used by Simon and Ericcson[9]) were issued to them. The examiner's role was to issue the practice verbalization problems and to answer any questions raised by subjects during the session. Upon issuance of the problem, the examiner proctored the sessions and reminded the subjects (if necessary) to resume talking if they failed to verbalize for more than three or four seconds. The sessions averaged approximately 40 minutes (including the practice verbalization problems), and tape recordings covered the extent of each session.

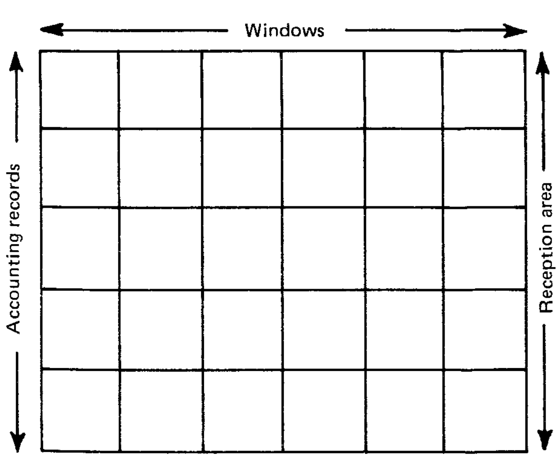
DATA PROCESSING, INSTRUMENTATION AND ANALYSIS

Pre-processing of protocol data

Pre-processing of the protocol data involved three steps: transcribing, segmenting, and encoding. Transcription of protocols was labour-intensive. However, a word-processing software was utilized to facilitate that process, as well as the subsequent stages of segmenting and encoding. Segmenting (or 'chunking') of the protocols was relatively simple, since verbal pauses were used in transcription to begin new lines of text. Other clues to chunking were hesitations and syntactically complete thoughts. The great majority of verbalizations were of short duration, consisting of only one line before a pause. In instances where a single line of data clearly indicated more than one discrete verbal behaviour, lines were broken into separate chunks.

Encoding of processed data

Encoding of protocol data was accomplished using CODEPRO, a computer program designed to control and lend increased validity and reliability to the process. CODEPRO is designed to flash up on the computer monitor individual chunks of protocol for evaluation and encoding. An individual (or encoder) then assigns one of nine codes (listed above) to each chunk. CODEPRO also allows encoders to return to the previous chunk (but not any farther) to reassign a code. Three paid encoders (one of which being the author) were used to assess the reliability of the instrument, or rather to determine the degree to which encoders agreed in their judgements. Each of the three encoders made over 2000 encoding judgements over the five protocols, and the average time taken to encode each protocol was approximately 75 minutes.



B uses the accounting records more than does D,
E meets people in the reception area more than does G,
A meets people in the reception area more than does C,
F uses the accounting records more than does E,
D meets people in the reception area more than does G,
C uses the accounting records more than does F.

A does not want an adjoining office with B,
F wants an adjoining office with A,
B does not want an adjoining office with G,
C wants an adjoining office with D,
E does not want an adjoining office with B,
G wants an adjoining office with F.

The total number of vertical corridors occupied by offices should be as small as possible.

F has more prestige than B,
C has less prestige than B,
G has less prestige than C,
D has more prestige than F,
E has more prestige than G,
A has less prestige than D.

Figure 2. The design problem

Analysis of the protocol data

A-Analysis based on encoding

Another computer program (CODESYN) was written for this study to assist in assimilating, tabling, and graphing the encoded data.

From the three separate encodings of the raw protocols (generated by the three encoders), an assimilated encoding was generated. The purpose of creating an assimilated encoding was to factor out some of the idiosyncrasies of encoding likely to be found in each of the separate encodings, and to base the assimilated version on some factor of inter-coder agreement. Therefore, in cases where three out of three, or two out of three encoders agreed on the nature of a chunk (which was the vast majority of cases) that chunk was automatically assigned the matched code. In cases where no agreement was noted between encoders, CODESYN assigned (defaulted) the code given by Encoder number 1 (the author).

CODESYN also assisted in calculating the degree of inter-coder agreement across the nine code variables (Table 1). Kuder-Richardson and Hoyt's reliability statistic (as reported by Kerlinger) was employed for the task. The degree of inter-coder agreement was high in five of the six analyses, including the overall analysis. The only protocol for which only moderate agreement was shown (0.70) was that for Subject 1.

From the assimilated protocols for the five subjects, CODESYN generated tables and graphs to chart the summed values of each coded verbal behaviour (in increments of 20 chunks) over the course of problem-solving. The graphs generated by CODEPRO are of particular interest since they appear to reveal the chronology of five unique design processes. From these tabled values, also a two-way analysis (ANOVA) was conducted to test the modest experimental hypothesis.

B-Analysis based on reference parsing

Two separate (and quite different) analyses of the protocols were conducted by parsing the raw protocols for occurrences of letter characters in the protocols (i.e. office workers: A, B, C, D, E, F, G), and central issues of the problem (i.e. Office Function, Prestige, Adjacency Preferences, and Compactness/Cost) (see Figure 2 – the problem). The rationale for parsing was to gain a rough idea of the degree to which subjects attended to central characters and issues of the problem.

Table 1. Inter-coder reliability coefficients for five protocols and overall

Subjects:	S1	S2e	S3e	S4	S5	Overall
Alpha Level:	0.70	0.96	0.98	0.97	0.96	0.98

Table 2. ANOVA of coded verbalizations for all subjects based on three encoders

Code:	LC	PC	IN	IP	MO	SE	SA	GA
S1	85	51	35	15	33	20	93	19
S2e	98	101	33	14	17	5	99	25
S3e	124	128	48	51	62	5	192	44
S4	59	67	70	9	23	18	59	11
S5	40	21	71	28	54	6	73	12

Source:	ss	df	ms	F	p
Total	68308.9	39			
Subjects	10374.65	4	2593.66	3.582	<.05*
Codes	37634.3	7	5376.33	7.425	<.001***
Error	20299.95	28	724.10		

RESULTS

The following results are based on: analysis of encodings, and analysis of parsed references.

Results of encoding analysis

ANOVA of coded verbalizations for all subjects based on three encoders (Table 2)

In this analysis, the assimilated encoding yielded the sums given in Table 2 for each code variable. Marked differences in verbal behaviour were noted both between and within the five problem-solvers. ANOVA confirmed these differences to be statistically significant. (Expert designers in the sample are designated by the letter 'e').

Results of parsing analyses

ANOVA of parsed verbal references to letter characters: A, B, C, D, E, F, G

This analysis (Table 3) involved the summing of letter character references parsed from the protocols. Marked differences in attention to letter characters (office workers) were noted both between and within the five problem-solvers. ANOVA confirmed these differences at levels that are statistically significant.

Table 3. ANOVA of parsed verbal references to letter characters

Workers:	A	B	C	D	E	F	G
S1	61	83	65	61	48	82	75
S2e	43	72	62	61	49	75	75
S3e	119	144	112	122	101	127	124
S4	41	73	48	55	61	56	64
S5	35	43	65	59	34	43	49

Source:	ss	df	ms	F	p
Total	28247.89	34			
Subjects	23764.19	4	5941.05	71.20	<.001***
L.Chtrs.	2481.09	6	413.52	4.96	<.01**
Error	2002.61	24	83.44		

ANOVA of parsed verbal references to central issues of the problem (i.e. Office Function, Prestige, Adjacency Preferences and Compactness/Cost)

This analysis (Table 4) involved the summing up of references to central problem issues parsed from the raw protocols. The analysis showed no significant difference between problem-solvers in their attention to central problem issues. However, a statistically significant difference was noted in the degree of attention paid generally by all to the central issues. Apparently the issue of least importance to most of the subjects was that of Compactness/Cost. Key words in the parsing were: 'accounting', 'reception' – 'adjoining', 'office', 'next to' – 'prestige', 'windows', 'window' – 'corridor', 'space', 'cost'.

Summary of results

Findings of the ANOVAs are generally consistent with prior expectations regarding the differential nature of problem-solving behaviour in designers. However, it was of some surprise (considering the small sample size) that the data indicate (Table 2) a significant statistical difference in the nature of information processed by subjects during problem-solving. It was not particularly surprising that some verbal behaviours (e.g. Literal Copies, Specific Assessments, etc.) were highly prevalent, whereas others were not. A balanced handling of information-types over the course of the design process was not expected, nor would it be desirable.

The remaining two ANOVAs (Table 3 – Parsed References to Characters and Table 4 – Parsed References to Central Issues) were merely informal analyses possessing questionable validity. However, what they reveal is of some interest. Table 2 indicates the variation in the problem-solvers' attention to letter characters. Subject 3e likely skewed the results (because the session was not a timed test) but even so, the degree of variance is interesting in regard to how differently the subjects attended to the various office workers.

Table 4 indicates that all the subjects established a

Table 4. ANOVA of parsed verbal references to central issues

*Issue:	1	2	3	4
S1	62	62	96	3
S2e	67	68	90	28
S3e	89	79	145	10
S4	56	27	99	15
S5	67	59	59	27

Source:	ss	df	ms	F	p
Total	23304.8	19			
Subjects	2491.8	4	622.95	1.91	NS
Issues	16900.	3	5633.33	17.28	<.01
Error	3913.	12	326.08		

*1, office function; 2, prestige; 3, adjacency preferences; 4, compactness/cost

hierarchy of problem issues. However, there was no significant difference between subjects in the nature of hierarchies.

The high degree of inter-coder reliability (Table 1) shown in the study indicates that the three encoders viewed the coding categories similarly, and that the matching of code to verbal behaviour was natural and unproblematic in most cases.

DISCUSSION

Because the present exploratory study was based a sample of only five subjects, the above ANOVA results are not particularly powerful (although they are a rich source for future experimental hypotheses). For that reason the following discussion will tend toward a case-study discussion of the individual protocols rather than elaborate comparisons of the protocols.

The results of the present study are most dramatically shown in graphs generated from the assimilated encoding (Figure 3). The graphs set the independent variable of Time (shown by increments of 20 chunks) against the dependent variable of Code Totals. The following is a

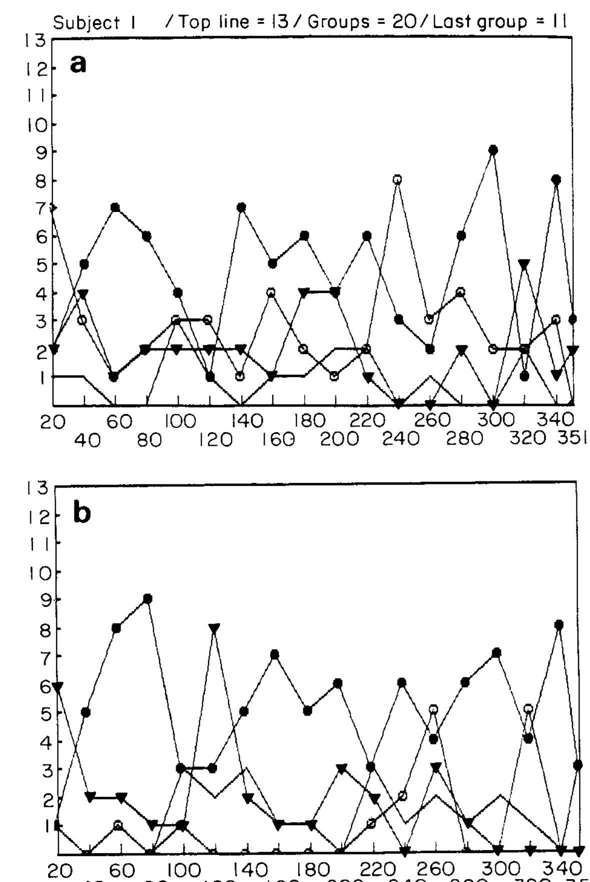


Figure 3. Graphs generated from the assimilated encoding: (a) Literal Copy (LC), ●; Paraphrased Copy (PC), ○; Inference (IN), ▼; Intention/Plan (IP), —. (b) Move (MO), ▼; Search (SE), —; Specific Assessment (SA), ●; General Assessment (GA), ○. (Graphs continue on next page)

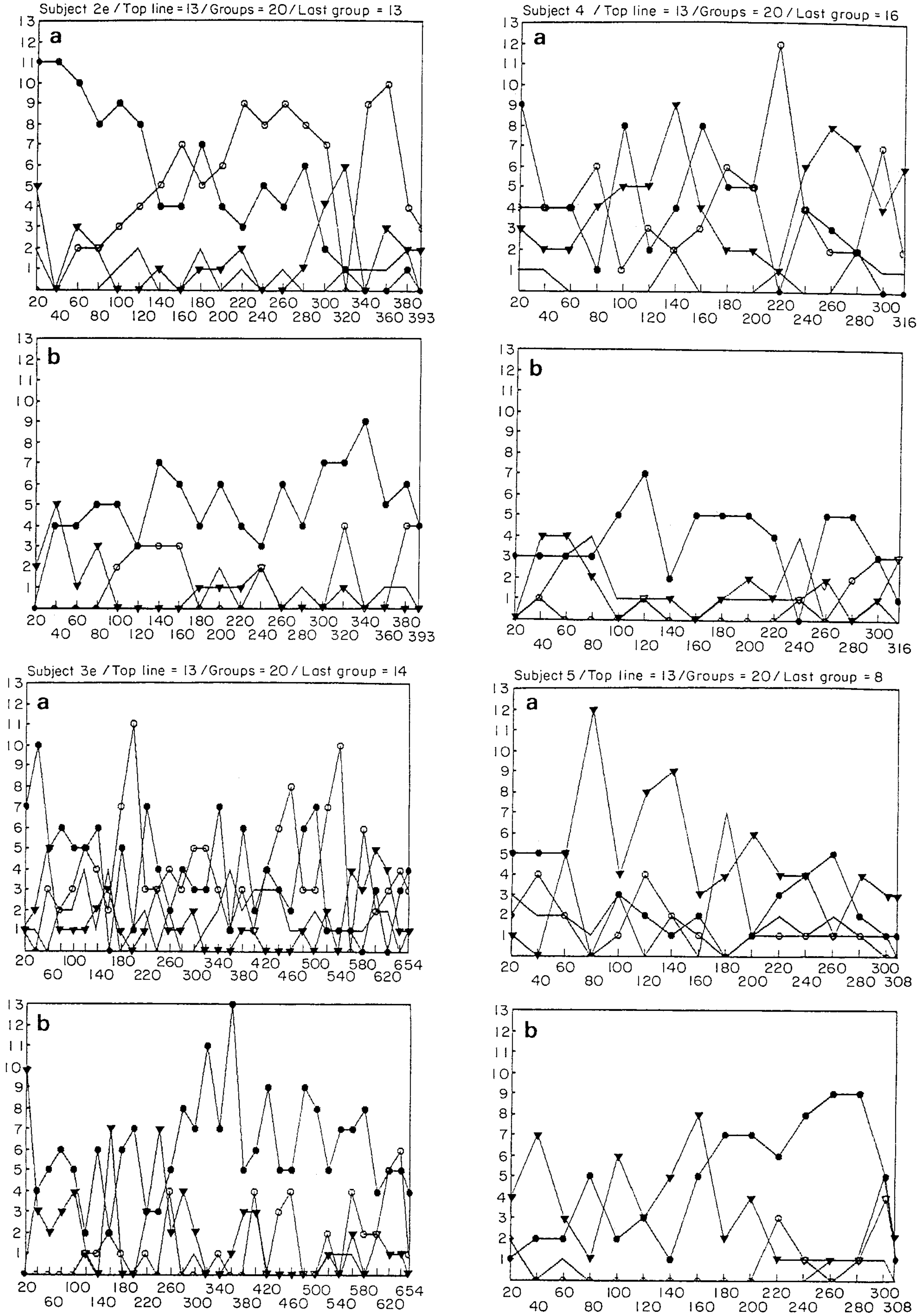


Figure 3 (continued): (a) Literal Copy (LC), ●; Paraphrased Copy (PC), ○; Inference (IN), ▼; Intention/Plan (IP), —. (b) Move (MO), ▼; Search (SE), —; Specific Assessment (SA), ●; General Assessment (GA), ○

more detailed interpretation of protocols for Subjects 1 and 3e – two protocols which illustrate quite different problem-solving behaviours. These interpretations are provided to show the manner in which encoded protocol data become useful information to the design scientist.

Subject 2e

Subject 2e was a practicing male interior designer, and he began his design process with much verbatim reading of the functional requirements (FRs). In fact, 32 of the first 60 encodings were Literal Copies (LCs). Only seldom (3 out of the first 60 chunks) did the subject phrase an FR in his own words. Interestingly, however, the lines for LC and PC show a strong interaction over the course of the problem, as LCs gradually decrease while PCs at the same time increase. This interaction appears to show how an internal representation of FRs develops over the course of problem-solving, and how problem understanding is gradually achieved. Inferences for Subject 2e appeared only sporadically during the process with no directional increase or decrease as was noted for LC and PC. This seems to indicate that inferential thought was not required constantly during the process, but was periodically, generated (sometimes in quantity) for the purposes of decision-making. Encodings for Intention/Plan (IP) behaviours were also sporadic throughout the process, but were less prevalent than Inferences – indicating that goals and strategies were generated as needed and not just in initial stages of problem-solving. Encodings for Moves (MO) and Searches (SE) were not particularly informative, and are being re-evaluated in light of these results. Of particular interest are encodings for the category of Specific Assessment (SA), as they indicate a directional increase in such thought from the start of the session to its completion. This directional increase resembles that of PC, and indicates a gradual increase in attention to specific organizational relationships in the office space. The percentage of SAs rivals those of LC and PC (each of which averaged approximately 25% of all encodings for Subject 2e). General Assessments (GA) were not as frequent as were Specific Assessments, however; neither did their graphed distribution resemble that of SA. An overall observation of these graph lines is that Subject 2e undertood a methodical approach to the problem, and patiently learned its parameters before fashioning its solution. The solution developed gradually through periodic decision-making, and to copious attention to its functional requirements as well as to the developing design itself.

Subject 1

Subject 1 was a female undergraduate interior design student at The University of Maryland. She began her design process by paraphrasing and copying the FRs about equally (17 of the first 40 chunks were copies – 7=LC, 10=PC). Over the course of the problem, however, Subject 1 showed much greater dependency on verbatim reading of FRs (LC averaged 24% of total encodings, with PC averaging only 14%) than on paraphrasing them. The charted lines for both LC and PC do not show directional increases or decreases, but rather sporadic peaks and valleys from beginning to end. This indicates, perhaps, the subject's failure to adequately internalize the FRs, and hence a failure to understand the essence of the problem itself. Encoded Inferences accounted for about 10% of total encodings, with a charted line showing moderate consistency throughout the course of problem-solving. The mode of six Inferences occurred near the end of the session after 320 chunks. Encodings for Intention/Plans were noted throughout the problem, but accounted for only 4% of all encodings. Move (MO) encodings showed considerable activity in the early stages of the session (accounting for 17% of the first 120 encoded chunks), but tailed off expectedly in the latter stages. Search (SE) behaviours present throughout the middle and final stages of the session indicate active referencing to FRs as a prerequisite to proceeding with the problem. Searches were usually noted immediately before LC and PC encodings, but never accounted for more than three encodings per 20 chunks. Specific Assessments (SA) occurred 93 times in the session of 351 chunks (26% of total), with a mode of nine early in the session. SA encodings show no apparent directional increase nor decrease, but reflect remarkably the percentage and distribution of LCs. These line similarities confirm the casual observation that SAs usually occurred immediately after Literal Copy encodings. General Assessments (GA) were found in just 5% of the total encodings, with a directional increase in the last 150 chunks. These data appear to show increasing attention in Subject 1 to office configurations involving more than three characters, with assessments of a more global nature. In summary, the encoded protocol for Subject 1 portrays an individual faced with a problem with no implicit method for solution. The solution path generated shows no operating method to progress incrementally toward solution of the problem. Rather, for every variable (with the possible exception of GA) Subject 1 exhibited erratic behaviour, and a general dissatisfaction with her own strategy.

GENERAL SUMMARY

This study was successful in testing a potentially powerful design research methodology. The working hypotheses were essentially confirmed, with some evidence to indicate that designers vary significantly in the nature and amount of information processed during problem-solving. The primary method of Protocol Analysis (with three independent encodings of protocols) showed remarkable differences in the design problem-solving behaviour of five individuals. The secondary methods (Letter-Character and Primary Issue Parsing) were helpful in assessing how the designers structured and prioritized information. These methods also give us a glimpse of how values (e.g. function over prestige, etc.) operate in the assignment of design priorities.

Much work remains to be done in order to establish Protocol Analysis as a valid design research tool. The potential for experimentally testing research hypotheses is great, provided larger subject-samples are utilized. For example, experiments could be designed to assess the degree to which design heuristics influence cognitive processes. Other PA methods could examine expert versus novice approaches to design problem-solving in a given domain. Educationally, PA could be of potential use as a diagnostic tool, or as an instrument to show students the 'form' of their problem-solving, of which they may be singularly unconscious. In the context of these potential applications of PA, utilization of the computer to facilitate the processing and analysis of PA data is a great encouragement.

REFERENCES

1 **Akin, O** An exploration of the design process in **Cross, N (Ed.)** *Development in Design Methodology* John Wiley & Sons, Chichester (1984)

2 **Simon, H A** *The sciences of the artificial* MIT Press, Massachusetts (1971)

3 **Thomas, J C, Lyon, D and Miller, L A** *Aids for problem solving* IBM Research Report RC 6468 (1977)

4 **Gagne, R M** *Problem solving and thinking. Annual Review of Psychology 10* Annual Reviews Inc, Palo Alto, CA (1959)

5 **Reitman, W** *Cognition and thought* John Wiley & Sons, New York, NY (1965)

6 **Eastman, C M** 'Cognitive processes and ill-defined problems: A case study from design' *Proc. Int. Joint Conference on Artificial Intelligence* Washington DC (1969) pp 675–699

7 **Thomas, J C and Carroll, J M** The psychological study of design. Design Studies **1** (1979) pp 5–11

8 **Simon, H A** 'The structure of ill-structured problems' *Artificial Intelligence* **4** (1973) pp 181–201

9 **Simon, H A and Ericcson, K A** *Protocol analysis: Verbal reports as data* MIT Press, Massachusetts (1984)

10 **Goor, A** *Problem solving processes of creative and non-creative students* (Doctoral dissertation, University of North Carolina). Dissertation Abstracts International 1974, **35**, 3517A (1974)

11 **Newell, A and Simon, H A** *Human problem solving* Prentice-Hall, Englewood Cliffs, NJ (1972)

12 **Miller, G A, Galanter, E and Pribram, K H** *Plans and the structure of behaviour* Holt, New York (1960)

13 **Ullman, D G, Stauffer, L A and Dietterich, T G** *Protocol analysis of mechanical engineering design* International Conference on Engineering Design, Boston, MA, USA (August 1987)

14 **Aaronson, A and Carroll, J M** *Intelligent help in a one-shot dialog: A protocol study* IBM Research Report RC 12667 (#55028)-Human-Computer Interaction (1986)

15 **Adelson, B and Soloway, E** *A cognitive model of software design* Report No. 342, Department of Computer Science, Yale University (1984)

16 **Caroll, J M, Thomas, J C and Malhotra, A** Presentation and representation in design problem solving *Br. J. Psychol.* **71** (1978) pp 143–153