Graphic design evaluation: towards a rule-based system

Thesis

How to cite:

© 1994 George L. Glaze

https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Version of Record

Link(s) to article on publisher’s website:
http://dx.doi.org/doi:10.21954/ou.ro.0000fb76

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
GRAPHIC DESIGN
EVALUATION

© George L. Glaze
GRAPHIC DESIGN EVALUATION
Towards a rule-based system

George L. Glaze
MA (RCA)

Thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Design Discipline
Faculty of Technology
The Open University
Milton Keynes,
MK7 6AA
U.K.

Author number: N17044675
Date of submission: 29 July 1994
Date of award: 28 April 1995

December 1994
ACKNOWLEDGEMENTS

The exercise and completion of this thesis was made possible with the assistance of Nigel Cross and Jeff Johnson whose supervision was instrumental in facilitating the programme of research from its inception. I would also like to thank Trevor Lambert who advised on the statistical approach and assisted with the statistical analysis. I am also indebted to the graphic design experts who gave of their time to provide the prescriptions used in the empirical study and the anonymous designers whose work provided the specimens used in the experiments.

I am grateful to the SERC and the Open University who provided the funding for the program; and Watford College who provided the initial industrial release. I wish to thank the staff of the Design Department at The Open University who provided access to equipment and references. And finally I would like to thank my wife for her emotional support throughout the proceedings and for her unstinting patience.
CONTENTS

ABSTRACT 6

CHAPTER 1: INTRODUCTION 7

  1.1 Motivation for the thesis  7
  1.2 Aim of the thesis  11
  1.3 Structure of the thesis  13

CHAPTER 2: GRAPHIC DESIGN 15

  2.1 Introduction  15
  2.2 The evolving graphic design idiom  18
  2.3 Variability in graphic design  19
  2.4 Underlying similarities with other design domains  20
  2.5 Models of design processes  22
    2.5.1 Environmental design  22
    2.5.2 Architectural design  24
    2.5.3 Mechanical engineering design  26
  2.6 Traditional scepticism to design systematisation  2
  2.7 Creativity and intuition  30
  2.8 The graphic design process: a case study  31
    2.8.1 Preparing an A4 publication using manual procedures  32
    2.8.2 Glaze's model of the graphic design process  45
    2.8.3 Tunnicliffe's computer assisted graphic design procedures  47
    2.8.4 Cotton's computer assisted graphic design procedures  49
  2.9 Similarities and differences in traditional and computer assisted graphic design process  53

CHAPTER 3: RULES IN GRAPHIC DESIGN 55

  3.1 Introduction  55
  3.2 Factors of typographic composition  57
    3.2.1 Distribution of elements in a page layout  59
    3.2.2 Formal structures and grids  60
3.2.3 Symmetry  74

3.3 Readability  76
   3.3.1 Psychology of perception  78
   3.3.2 Legibility  79
   3.3.3 Column width and optimum line length  81
   3.3.4 Variety in type faces and type sizes  83
   3.3.5 Capital letters  84
   3.3.6 Type setting  85

3.4 Production constraints  86

3.5 Can we define ‘good graphic design’?  87

3.6 Can we define ‘bad graphic design’?  90

3.7 We can define ‘not bad’ designs  92

3.8 Conclusion  92

CHAPTER 4: RULE-BASED SYSTEMS  95

4.1 Introduction  95

4.2 The architecture of a rule-based system  96
   4.2.1 Triggering and firing  100
   4.2.2 Conflict resolution  100
   4.2.3 Control mechanisms  101
   4.2.4 Forward chaining  102
   4.2.5 Backward chaining  103
   4.2.6 Dialogue module / human computer interface or user interface  104

4.3 Advantages and disadvantages of rule-based systems  108

4.4 Difficulties in using rule-based systems  109

4.5 Use of rule-based systems for design evaluation  110

4.6 Arguments for and against the validity of expert systems  112

CHAPTER 5: RULES DERIVED FROM A PRIOR SOURCE  115

5.1 Introduction  115

5.2 Knowledge acquisition and elicitation  115

5.3 Knowledge elicitation for graphic design  118

5.4 Rules extracted from a published source  120
5.5 Rubinstein’s recommendations 121
   5.5.1 Formulating rules from Rubinstein’s recommendations 122
   5.5.2 Can Rubinstein’s rules be applied to books as intended? 123

5.6 Summary 125

CHAPTER 6: THE EXPERT SYSTEM TEST 126

6.1 Introduction 126
6.2 The formulation of representational rules from natural language 126
   6.2.1 First iteration 127
   6.2.2 Second iteration 128
   6.2.3 Can the Rubinstein recommendations be applied to other populations of publications? 129
   6.2.4 Can the Rubinstein rules be applied to desktop publications of A4 posters and notices? 132
   6.2.5 Experimental shortcomings and revisions 134
   6.2.6 Can difficulties and ambiguities be removed? 135
   6.2.7 Third iteration 135
   6.2.8 Can these rules be used by others? 137
   6.2.9 Fourth iteration: Testing the formulated rules using the Leonardo expert system 144

6.3 Conclusions 152

CHAPTER 7: EMPIRICAL STUDIES - RULES ELICITED FROM GRAPHIC DESIGN EXPERTS 153

7.1 Introduction 153
7.2 Rules elicited from graphic design experts? 156
   7.2.1 Preparation of the sample of forty posters and notices 157
   7.2.2 Description by the experts of the ‘bad symptoms’ detected in each specimen of the sample 159
   7.2.3 Interpretation of free format speech to form technically defined and rigorous rules 160

4
7.3 Testing the rules elicited from the designers 174
   7.3.1 Experimental design and method of analysis 174
   7.3.2 Application of the rule sets 175
   7.3.3 The experts' evaluation of the posters 178

7.4 Analysis of the data - visit 1 179

7.5 Follow-up interviews 188
   7.5.1 Conclusions of the follow-up interviews 188
   7.5.2 A follow-up test to see if expert A uses his rules consistently after ten months 191
   7.5.3 Some possible causes of the experts' inconsistency 194
   7.5.4 Summary 196

7.6 Conclusions & discussion 198

CHAPTER 8: CONCLUSIONS 209
   8.1 Introduction 209
   8.2 Summary of research findings 209

CHAPTER 9: FURTHER WORK: 211
   9.1 Introduction 211
   9.4 Further researches 212

REFERENCES: 217

FIGURES AND TABLES: 230
ABSTRACT

Current trends in desktop publishing (DTP) systems enable novices to prepare a variety of desktop publications. These systems provide limited graphic design support to users inexperienced in the graphic design domain. While to varying degrees they facilitate the deployment of text and image components by grid oriented and column based prescriptions, they do not address the universe of aesthetic issues. It is suggested that automatic procedures can provide reliable systematic evaluation of the aesthetic quality in a desktop publication for this group of users.

This study describes a novel procedure for the elicitation of aesthetic rules from graphic design experts. Experts were asked to diagnose bad symptoms in DTP layouts, and rules were derived from their diagnoses. The diagnosis of 'bad' design symptoms in a graphic layout is shown to be an effective method of making explicit the graphic designer's implicit, intuitive aesthetic judgements. A rudimentary but systematic procedure for the application of graphic design rules by non-designers, simulating expert graphic design evaluation, is reported. The preliminary design of a prototype rules-based system devised to complement the novice's design abilities and performance in the aesthetic evaluation of a graphic layout is described.

The results of the empirical tests show that some graphic designers employ intuitive, implicit aesthetic rules in the evaluation of a graphic layout and these can be made explicit; that graphic design rules can be applied systematically to a population of graphic layouts; expert graphic designers have rules, but they apply them inconsistently; and graphic design aesthetic expertise can be elicited and entered into a computer system to enact the role of a graphic designer. The performance of a systematic application of a graphic designer's rules is shown to be more consistent in design evaluation than the graphic designers from whom they were elicited.

It is concluded that a method has been established which allows some aspects of graphic design aesthetic evaluation to be systematised.
CHAPTER 1: INTRODUCTION

1.1 Motivation for the thesis

The availability of sophisticated desktop publishing systems provides novices and experts alike with the ability to prepare a variety of desktop publications. However, there exists some scepticism by some graphic designers and researchers about the aesthetics of some of the graphic designs produced.

The graphic designer’s principal contention is that the aesthetics of a graphic layout relies on the knowledge of well established principles and observance of their application. They speculate that novices who are unaware of these principles are unable to achieve good design in the production of a desktop publication.

Some researchers are in agreement with the graphic designers’ view, and therefore stress the need for desktop publishing systems to incorporate the rules of graphic design via an expert system. ‘One of the most promising developments in the use of computers in recent years has been the work on expert systems. Its significance is that it addresses itself to providing computer systems which are able to make a “knowledgable” contribution to complex problems in a specific domain or field of interest, that is, to act as an “expert”.’ MacCullum and Duffy (1987). Nevertheless it is doubtful that many graphic designers would support this proposal.
However, according to researchers like Casner (1991); Baker (1988); Overmier (1988); Theophano (1992); Corbitt (1991), there is need for effective tools for design evaluation tasks in the desktop publishing domain but they acknowledge the difficulty of the proposition. Typical of these researchers are North and Johnson (1990) who report that 'The aesthetic appeal of a business publication can be enhanced by applying a few basic ingredients in the right proportion and at the proper location within a document'. They suggest that 'The four most universally accepted criteria of page design are (1) balance, (2) consistency, (3) contrast, and (4) focus.' But these researchers agree that graphic design rules are difficult to prescribe, relying on professional graphic design expertise; and as Blair (1989) points out 'DTP requires more than a little knowledge of both computing and publishing and the skill to tread a fine line between the two'.

Nevertheless, it is perhaps possible that some of the graphic designer's elementary knowledge of graphic layout can be extracted from the rudimentary recommendations reported in the literature on graphic design. So, a simplistic solution to this problem would be initially to provide users of desktop publishing systems with access to the appropriate graphic design recommendations and guidance on the routine ways these recommendations should be applied.

To discover whether DTP, word processing and typesetting systems contained graphic design rules I undertook a survey of the leading software magazines, eventually contacting major suppliers and various producers. Table 1.1. gives
Table 1.1: Rules and options in major DTP and word processing software.
the results of information provided by these sources. In several instances the information was correlated by reference to manuals and demos of software packages provided by suppliers and producers. For LATEX a logical design system, the data is augmented by information obtained from the literature. It was found that some producers of visual design systems attempt to provide what some describe as templates and others call AutoFormat or style sheet tools which enable the user to choose a professional document design for a publication. According to Howard (1994), of pre-sales support at Quark, their product is ‘aimed at the professional graphic design market’ where the graphic designer’s wish for individuality can be optimised.

Of the word processing packages Word 6 is arguably the most ambitious. Kirby (1994), of technical support at Microsoft says Word has no graphic design rules but contains over 70 user selectable style sheets ‘so the user can view a document in a variety of formats before choosing the most suitable one’.

Programmable typesetting systems such as LATEX contain well established logical design rules of typesetting, see Lamport (1986). TEX the program on which LATEX is based, incorporates typographic rules for:

- widows and orphans i.e. so a single word will not appear at the top of a page, instead other text will be brought with it.
- justification i.e. justification will be used except in the last line in a paragraph where proportional spacing will be used.

Lamport distinguishes LATEX as encouraging sound typography as opposed to visual display systems which he suggests mistakenly rely on aesthetics alone.
However Boag (1994), of Monotype dismisses $\TeX$ and $\LaTeX$ as being for mathematicians, programmers and scientists. According to Boag this system is wholly unsuitable for the average DTP user; it is not a mainstream DTP package.

Radcliffe (1994), the product marketing manager for product publishing for Aldus confirmed that 'other than defaults for margins our systems contain no rules as such', because in her view, 'rules would impede the flexibility which designers prize'. Heath (1994) the marketing director of Apple UK, observes that word processing and DTP packages do not contain rules but provide an increasing variety of user options to enable the user to incorporate their own rules. But providing every thinkable useful feature is confusing for novices and inhibits learning. There is disagreement between the producers of visual and logical design systems. There is also confusion in the graphic designer's application of some of their rules. So the difficulty is in revealing the graphic designer's implicit knowledge of graphic layout aesthetics for these systems.

1.2 Aim of the thesis

The principal aim of the research reported in this thesis is to investigate the possibility of providing an automated rule-based system of graphic design evaluation. Such a system could conceivably be implemented as part of, or an addition to, a desk-top publishing system, in order to give inexperienced designers some form of evaluation of the page layouts they produce. The aim of the research reported here was not to develop fully such a system (although a rudimentary version of such a system was developed as part of the research).
Rather, the aim of the study has been to find ways of answering the following questions; firstly, do graphic designers employ intuitive, implicit aesthetic rules in the preparation of a graphic layout? Secondly, if they do, can their aesthetic rules be elicited? And thirdly can the graphic designer’s aesthetic expertise be entered into a computer system to enact the role of a graphic design expert?

To address these issues it would be necessary to pursue a variety of secondary questions. For example, what is graphic design? What is good graphic design? Can the graphic designer’s intuitive rules be made explicit? At what stage in the design process are aesthetic rules applied? Do graphic designer’s have their own personal aesthetic prescriptions? Are the graphic designers prescriptions influenced by tradition? Do graphic layouts differ from other forms of layouts? What conditions the (non)application of graphic design rules? To what extent are graphic design rules for one population of graphic layouts generalizable to other populations? Do graphic designers apply their rules consistently? Can graphic design rules be entered into a computer system? Can graphic design rules be applied systematically to a population of graphic layouts?

The overall purpose has been to enhance our understanding of the expertise involved in graphic design evaluation, using the methodology of rule-based systems. Knowledge elicitation requires the knowledge expert to articulate their expertise using a variety of well established scientific procedures. The
difficulty is that most artists and designers find it difficult to articulate fully their rationale. (Cross, A. 1984). Therefore a novel process of eliciting the required knowledge from graphic design experts is required.

1.3 Structure of the thesis

This thesis is divided into a further eight chapters. The theme in chapter 2 is the graphic design process. The genealogy of typography is traced from calligraphy to computer typesetting. Graphic design influences and tradition are reported. The nature of graphic design is examined.

In chapter 3 the principles which affect the product of graphic design are examined. This aspect of the study examines the literature concerned with the canons associated with the elements in a graphic layout which contribute to its wellformedness. Among these are pertinent recommendations drawn from legibility researches and anecdotal accounts given by acknowledged graphic design authorities. A new approach to graphic design evaluation is given based on the proposition that it is possible to diagnose and report the presence of design defects (bad design) in a sample and that this diagnosis can be used to identify the same defects in other samples in the same population. The corollary is that the absence of design defects is an indicator that the design representation is not bad. It is argued that the diagnosis of bad design symptoms can inform design evaluation and indeed is a more rigorous instrument for design evaluation than one which seeks the unequivocal declaration of good design.
Chapter 4 discusses rule-based systems. It describes the architecture of an expert system and examines the utility of rule-based systems for graphic design evaluation. The opinions of proponents and opponents of expert systems are discussed.

In chapter 5 procedures for the acquisition of graphic design rules are described; in particular the elicitation of graphic design knowledge from recommendations available in a published source is examined.

Chapter 6 reports the Expert System Test (EST); a rigorous process used to remove interpretational difficulties in the elicitation of rules from complex systems like graphic design and to test the formulation and validity of the rules devised.

In Chapter 7 the focus is on the empirical study of a procedure to elicit intuitive aesthetic rules of graphic layout design from two graphic design experts. The effectiveness, predictability and reliability of rules elicited from a graphic design expert and the feasibility of their systematization is examined. A simple procedure for human inspection and diagnosis designed to simulate an interactive rule-based process for A4 layout evaluation diagnostics is reported.

Chapter 8 gives a list of the conclusions reached in this study. Finally, chapter 9 attempts to forecast developments of desktop publishing systems which can provide design evaluation of desktop publications.
CHAPTER 2: GRAPHIC DESIGN

This chapter provides a background to the graphic design domain through a survey of the literature on graphic design, architecture and engineering design methodology. It examines the evolution of the graphic designer's conventions, terminology and procedures. It reports the similarities and differences of design methodology and compares a traditional graphic design model with a contemporary desktop publishing model.

2.1 Introduction

Muller-Brockmann (1971) attributes the origin of graphic design to a form of visual communication which can be traced to a deliberate form of mark making which emerged from primitive cave painting, ideograms and knot writing. The latter remains in use today in some middle eastern communities. According to Meggs (1983) 'It was not until 1922, when the outstanding book designer William Addison Dwiggins coined the term “Graphic Designer” to describe his activities as an individual who brought structural order and visual form to printed communications, that an emerging profession received an appropriate name'. Rand (1985) explains that 'the term was used by W.A. Dwiggins in a statement which first appeared in the Boston Transport, August 29, 1922, as follows: “Advertising design is the only form of graphic design that gets home to everybody”. Whether Dwiggins is the original source of this appellation is a matter of conjecture'.
Until Dwiggins, the term commercial artist was associated with practitioners of the disciplines which characterise graphic design; publicity design, advertising design, illustration and book design, etc. But according to Rand (1985), snobbery as well as the lack of clarity about its scope may have contributed to the term's gradual disappearance from the graphic arts scene today. The term 'graphic design' has since become a ubiquitous name for the work of designers who contribute to advertising design as well as those engaged in other forms of visual communications such as book design.

Lewis and Brinkley (1954) report that Professor Richard Guyatt's adoption of the name Graphic Design for his department at the Royal College of Art in 1950 helped to confer academic approval on Dwiggins' nomenclature.

According to Rand (1985) 'the expression graphic design is rich in meaning but difficult to pin down'. He reservedly gives this definition of a graphic designer; 'Specifically, a graphic designer is one who creates ideas that are expressed in words and / or pictures, and generally solves problems of visual communication'. The term graphic design therefore encompasses a spectrum of activities which rely on visual representations; all forms of fine art, illustration, photography and animation, belong to the genre of visual communication.

Twyman (1970) relates the ancestry of graphic design more closely to the crafts and traditions of the burgeoning book printing industry of the eighteenth
century. He confines graphic design to the processes concerned with the orchestration of type and non-type elements on a page or poster layout, focussing on the typographic reliance of graphic design.

There is some confusion over the use of the terms 'graphic design' and 'typography'. In Mclean's (1980) view graphic design is a synonym for typography. He reports that the roots of typography are in the calligraphy of 'an eastern corner of the Mediterranean less than two thousand years before Christ'. Typography, he affirms, is a mark making activity which has successfully adapted to changing communication needs and technological advances. Zachrisson (1965) defines typography as 'the art of selecting and using type faces and typographic material with the aim of transmitting a message, ultimately by means of printing.' Lamport (1986) characterises typography as a craft concerned with 'logical design' and with 'making a document easier to read'. In his view, a typographer is concerned with 'logical structure, not visual appearance'.

Tschichold a seminal book designer, typographer and graphic designer offers the following distinction between graphic designer and typographer (Tschichold, 1991). He contends that graphic design is concerned with self expression while typography is the province of designers with a desire to conform with typographic traditions and conventions.

Rubinstein (1988) observed that 'digital typography includes computer-aided
design of books and other documents, tasks traditionally performed by graphic artists, typographers and book designers’. In his view graphic artists are different to typographers but he does not elaborate on how they may be distinguished.

It is concluded that a typographer is a graphic designer who is concerned for the integrity of the type at his command. Good typography should be the concern of all graphic designers. Therefore, an experienced graphic designer when incorporating type into his designs is exercising the skills of typography. Graphic design and typography are complementary, interdependent activities; macro and micro levels of the same design domain. Graphic design is a highly specialised activity which is principally concerned with achieving the successful communication of visual messages. In all graphic designing the priority is to ensure appropriate interpretation of a visual representation. Graphic design is employed in a diverse and rapidly growing industry which includes multimedia (e.g. TV graphics, computer graphics) as well as the more conventional paper based, printing related domain.

2.2 The evolving graphic design idiom

Over time graphic design, like many other disciplines, has evolved its own highly specialised procedures and language. Graphic designers have willingly, if not always initially enthusiastically, embraced related technological developments. Technological advances often provide opportunities for innovation. Many of the graphic designer’s processes are reliant on print
quality and visual reproduction techniques to enhance the message transmission. This has necessitated the adoption of associated terminology.

Increasingly the graphic designer's processes have expanded to incorporate a variety of novel technical procedures. Recent technological advances have resulted in the introduction of many new terms into the designer's vocabulary and idiom. 'Fount' for example, traditionally describes a set of alphanumeric characters in one design and of one size. In desktop publishing applications a font means a typeface in a complete range of styles and size variants; both terms are used in this text. The format of a page traditionally described as being either portrait or landscape is now described as being tall or wide. Type 'alignment' is used instead of 'typesetting'. Consequently, the context of particular terms has to be declared, or understood beforehand in order to avoid confusion in their interpretation as the repertoire of graphic design processes is expanded.

2.3 Variability in graphic design

Variability in the graphic designers procedures results from a combination of circumstances. Every graphic design activity is different due to the variety of contents and prevailing production constraints. The same task repeated by an individual designer will be addressed differently. A task undertaken by one individual will be executed differently when undertaken by two individuals working together. Different tasks undertaken by the same designer will be executed differently as a consequence of changing needs. A task repeated by
a designer will be done differently due to the accumulated experience. A design task will be done differently in different environments; at a design studio and at home, for example.

The availability of technical resources will bring about a change in the design procedure. For example a desktop publishing system connected to a laser printer can provide an acceptable form of typesetting for general use. This facility gives the designer control of the typesetting process. However access to computer graphics facilities for example, can be useful to the experienced user but frustrating and time absorbing to the novice. Depending on the level of available technical expertise, technical resources may or may not improve the graphic designer's productivity.

2.4 Underlying similarities with other design domains
Like some, but not all, other forms of designing, graphic design is essentially concerned with solving visual problems. Graphic design differs from other forms of designing in three ways. Firstly, it is concerned with selecting visual elements. Secondly, it is concerned with the deployment of text and non-text elements to achieve precision, legibility and readability. Thirdly it is concerned with the configuration of visual elements to achieve visual impact, emphasis, sequence and order. Many of the graphic designer's traditional canons are shared with painters, architects and engineers. Common rules of configuration, structure, symmetry and asymmetry are observed in these activities. For example, one of the most common beliefs is that a surface can be divided into aesthetic proportions following principles laid down by the Ancient Greeks in a formula described as the 'Golden Section'.
There is disagreement on the value of simplification. It is often stated that simplicity of form in a graphic design solution is preferred to visual complexity, e.g. Brochmann (1970). Tufte (1990) argues that 'clutter and confusion are failures of design, not attributes of information'. He also reports Robert Venturi’s comments on architecture that blatant simplification means bland architecture: 'Less is a bore'.

Graphic designers share many similarities with other designers. Some of these may be witnessed during the analytic phases and where pencil and paper are the tools used for visual expression. Others are less evident but are apparent in the nature of the problems designers encounter and the strategies they employ to resolve them. For instance Cross (1985) illustrates dichotomies of cognitive style as:

- Divergent/Convergent
- Impulsive/Reflective
- Field-independent/Field-dependent
- Holistic/Serialistic
- Focused/Flexible

He stresses that different cognitive styles may be called upon throughout the design process. It is suggested that designers employ changing styles throughout the design process; moving from the Impulsive and Divergent stage where the designer initially considers the design options to the Reflective and
Convergent stages. For example, Lawson (1980) suggests that ‘what makes
design such a challenging task psychologically is the very even balance of the
two sets of mental skills that are needed to produce creative work.’ But we
don’t know if this applies to graphic design.

2.5 Models of design processes
A recent study by Tunnicliffe (1991) suggests that there is a paucity of
knowledge on the graphic design process. However, there is a substantial
body of opinion concerning the activities of the engineering and architectural
design processes.

Design is a creative genre in which the behaviour exhibited in its practitioners
reveals certain observable similarities in the practices and techniques
employed. Graphic designers employ observable generative procedures
similar to engineering and architectural designers. They are thought to
employ problem solving strategies like other designers (Schenk 1991).
Evidently many researchers have attempted to formulate the mental processes
of engineers and architects. Here some pertinent examples from research in
environmental design, architecture and engineering are reported to inform our
understanding of the graphic design process.

2.5.1 Environmental design
Halprin (1981) stresses the variability, interdependence and outward looking
nature of design activities. In Halprin’s model (see Figure 2.1), which he
calls the "RSVP cycle", he identifies the four main ingredients as:

**R** - Resources which are what you have to work with. These include physical resources, motivation and aims.

**S** - Scores describe the processes.

**V** - Valuation which analyses the results of action, possible selectivity and decisions; action oriented and decision-oriented aspects.

**P** - Performance which is the resultant of scores and the style of the process.

Scores are defined as processes which extend over time. We are perhaps more familiar with 'musical scores' says Halprin, but he extends scores to mean anything 'from a grocery list to a calendar'. He stresses that his model is equally valid in any direction as each activity can feed forwards or reverse. Moreover, he reminds his readers that while 'scoring makes the process visible' it is 'not meant to categorize or organize but to free and unblock' the process. In Halprin's words 'scores are not orderly, nor do they attempt to makes things function well. Scores are exploratory and not finite. Scores are open not closed'.

![Halprin's model of the environmental design process (Halprin 1981)](image)

Figure 2.1: Halprin's model of the environmental design process (Halprin 1981)
Consequently it could be argued that while the strength of RSVP is its flexibility, its weakness is its descriptive pattern which may make a particular scenario difficult to formulate and evaluate. RSVP is a cleverly devised model of a process which is multi directional and intersectional. This model is far too wide ranging to be suitable for the representation of graphic design which is essentially a serial process. As will be shown later in this chapter, graphic design features a series of progressions with few iterative activities beyond the conceptual design phase.

2.5.2 Architectural design

March’s (1976) model also features the cyclical or iterative nature of the design process, see Figure 2.2. This model has three distinct components - the first stage being Productive, the second Deductive and the third Inductive. In the Productive phase a design proposal is formulated after preliminary solution types have been identified from the initial requirements. The designer is then able to ‘deductively’ analyse the performance characteristics and to ‘inductively’ evaluate the possibilities of making further improvements in the design. March introduces the philosophical work of C. S. Peirce to explain the notion of synthesis or Abduction as central to design activity. March quotes, Peirce suggesting that ‘Deduction proves that something must be; Induction shows that something actually is operative; Abduction merely suggests that something may be’.
Figure 2.2: March's model of the architectural design process (March, 1976)
2.5.3 Mechanical engineering design

French's (1985) model of the engineering design process (see Figure 2.3), progresses through the following stages:

Analysis of the problem.

Statement of problem.

Conceptual design - where the possibilities are considered.

Selected schemes.

Embodiment of schemes - development of concept.

Detailing - redesigning particular features after tests.

Working drawings etc.

Feedback loops interconnecting analysis, conceptual design and embodiment of schemes characterize these important procedures. Analysis is described as 'a small but important part of the overall process' producing a statement of the problem. From a statement of the problem the designer is able to generate broad solutions in the conceptual design phase.

In the phase described as embodiment of schemes broad solutions are developed, evaluated and fed back to the conceptual stage if necessary. Detailing involves the designer in making the remaining important decisions.

This model is a closer representation to the graphic design process. However, unlike engineering, detailing is not essential to all forms of graphic representation.
Figure 2.3: French's model of the mechanical engineering design process (Cross (1989))
Pahl and Beitz (1984) divide the design activity into four main operations: clarification of the task, conceptual design, embodiment design and detail design. They emphasise the iterative nature of the activities in the design process and observe that the stages do not follow rigidly one after the other, see Figure 2.4.

Notably, Halprin considers the widest associations to be pertinent, unlike French and March who for example do not appear to consider the client's involvement to be relevant to the design process. This may relate to the status of the client in the various design methodologies.

2.6 Traditional scepticism to design systematisation

Many authors have reported on the scepticism of the designer on propositions relating to the systematizing of their processes. However, Lansdown (1982) points out that the prevailing evidence supports the view that designers often fail to apply relevant information and established techniques because of the lack of a systematic rigor. However, graphic design, unlike some other forms of designing, has little scientific reliance. Unlike architecture and engineering, there are few explicit absolutes in graphic design. Many creative professionals and academics question the relevance of design methodology for designers but accept its importance as a diagnostic support for technically demanding procedures. For example, Potter (1980) is critical of the arguably 'necessary simplification' and terminology employed in design methodology and sceptical of systems and creative prescriptions.
Figure 2.4: Pahl and Beitz's model of the engineering design process (Pahl and Beitz, 1984)
2.7 Creativity and intuition

According to Archer (1965) creativity is central to the design process. In his opinion 'One of the special features of the process of designing is that the analytical phase with which it begins requires objective observation and inductive reasoning, while the creative phase at the heart of it requires involvement, subjective judgement, and deductive reasoning. Once the crucial decisions are made, the design process continues with the execution of working drawings, schedules, etc. again in an objective and descriptive mood. The design process is thus a creative sandwich. The bread of objective and systematic analysis may be thick or thin, but the creative act is always there in the middle'.

Creativity in design has been variously described as being the consequence of brain-storming, evolving heuristics or using one's imagination to conjure up a wide variety of loosely connected initial thoughts broadly related to the design problem in an attempt at finding relationships. Designers believe that brain-storming reveals conscious or subconscious connections which emerge to be discussed with ones colleagues, for joint reaction and speculation, written down or sketched for later appraisal, review, evaluation and development.

Much has been said and written about the creative process. Mooney and Razik (1967), for example, report that 'creative behaviour, by its very nature, is spontaneous, inner-directed, ordinarily not capable of being elicited at will. Therefore, it is unpredictable and escapes manipulation and control. It is generally not amenable to experimentation'. Cross (1982) and Simon (1973)
suggest that design is a class of ill-defined problem. Tunnicliffe (1990) describes an ill-defined problem as 'one in which the requirements, as given, do not contain sufficient information to enable the designer to arrive at a means of meeting those requirements simply by transforming, optimising, or superimposing the given information alone.' As Cross (1982) comments: 'the solution is not simply lying there among the data, like the dog among the spots in the well known perceptual puzzle; it has to be actively constructed by the designer's own efforts.' Pye (1978) emphasizes that 'design, of every kind, is a matter of trial and error.' Creative thinking tends to result from concentrated, perhaps obsessional periods of convergent and divergent thinking according to de Bono (1970).

Little has been written specifically about design intuition. Intuitive thinking is reported to be a dynamic element of the creative endeavour. According to Hurlburt (1977) the graphic designer intuitively determines the proportions and format of a layout as a result of the prerequisite analysis of the contents, editorial objectives and their broad approach to the composition of visual elements to form a layout. But this ignores the constraints on creativity; graphic designers need to consider the prevailing production constraints of page size, format and sequence which contribute to the design outcome and often follow established rules of composition.

2.8 The graphic design process: a case study

We need clarification of the procedures undertaken by graphic designers. The
following describes a case study of the graphic design procedures involved in the design of a leaflet. The data obtained inform the preparation of the graphic design process. The subsequent impact on graphic designers of using desktop publishing systems is presented.

Potter (1980) wrote ‘graphic designers very often begin with a false brief from their client which will need to be taken right back to its origins before design (in the ordinary sense) can begin.’ Twyman (1970) suggests that when confronted by a new problem, graphic designers should first ask the question, ‘how can what we want to communicate be ordered so as to be most easily understood?’

Hurlburt (1981) likens the graphic design activity to ‘running a maze. The designer selects a line to follow only to learn that the constraints he encounters send him back to probe another direction until he finds a clear path to the solution. In the final analysis each designer will have to work out his own procedure to suit his thought process, working methods, and personal style.’

2.8.1 Preparing an A4 publication using manual procedures
The following study reports the procedures involved in the preparation of an A4 size publicity leaflet. This study is based on a live commercial graphic design commission which I undertook in 1991. It is being used here as an example of the graphic design process.
(i) **Background:**

The client is a sales executive of the Tokyo Electric Company (TEC). The company is involved in the manufacture, distribution and sales of electronic weighing machines, cash registers and barcoding scanners to the retail industry: they are the market leaders worldwide and aim to be market leaders in the UK. Their UK head office is in Watford. The company has a network of premises across the UK from which they produce, distribute, provide training and after sales support. They have a track record for quality, reliability and innovation. The designer has had a long association with the company for which he has completed several design commissions; he was therefore familiar with the company philosophy and their aspirations. The client (sales executive) is new to the company and this was their first meeting.

(ii) **Client Briefing:**

The company had developed a new product based on a system for producing barcode labels. The unique selling proposition (USP) was that any small business equipped with their software and thermal printer could now produce barcode labels incorporating their symbol or logo for themselves reducing the time and expense incurred by having to refer to a specialist supplier. The promotion of the product was to be low key and a variety of below-the-line advertisements to be placed in trade journals would be based on the design of a publicity leaflet.

The client was referred to the designer by the sales director with whom he was
previously acquainted. The designer was self-introduced to the client by telephone prior to a personal meeting. The client's brief was verbally delivered in the informal setting of the R&D workshop adjacent to the sales executive's office.

During the correspondence the merchandise, barcode system and method of production were described by the client; important details and technical specifications were noted by the designer. A skilled systems operator and a technician were introduced and a demonstration of the equipment was given. New and specialist terms were used by these personnel and the client for which clarification was requested and received.

(iii) Establishing the client's needs:
A good rapport was established in the interaction; design briefing owes much to listening, good interpersonal relations and mutual respect being quickly established. The client produced for discussion several publications he had collected from other manufacturers; one leaflet was from a major competitor with which the client was evidently impressed. The designer encouraged the client to reveal his likes and dislikes by identifying particular features and details on these examples; the designer used the opportunity to voice his own views. It was important to elicit the client's prejudices and rationale so that these could be avoided in the design solution. The personal briefing presented the opportunity to discuss the essentials of the publicity leaflet, e.g. understanding the message to be conveyed and the development of an audience
profile. The client explained that the leaflet was intended for the existing network of wholesalers of their products and to invite new enquiries directly from retail outlets. The wholesaler relied on technical and performance data while the retailer was persuaded by visual appearance, convenience and reliability.

Some copy had already been written and this was scrutinised by the designer whose comments were requested by the client. The designer advised on corrections to be made for the next meeting. Copies of the material discussed were made for the designer's retention. Additionally both the designer and the client spontaneously made a variety of thumb-nail sketches for record and clarification. The likely pictorial content was discussed and photographs and illustrations were considered. The designer recommended a photographer for additional photographic sessions.

The company has an established house-style from which the possibility for deviation was examined. As the correspondence progressed the client revealed that he would like his brother to produce the artwork (meaning illustrations) and examples of his work were requested by the designer for a follow-up meeting arranged for the following week. The specification, budget and completion date were agreed before the meeting closed.

(iv) Assimilation:
Notes made and the information gathered were examined on return to the
design studio. The commission was discussed with colleagues who gave comments and offered advice on alternative stylistic approaches, narrative and configuration. The client's message and the expectations of the anticipated wholesale / retail audience were also examined.

(v) Analysis:
The information was appraised and an estimate of the copy (heading, subheading and text) was made. The production constraints were assessed in terms of house style, the number of printings (colours to be used), etc. Paper size, margins, columns, type size and type style were also examined. Production of the publicity leaflet relies on the collaboration of a variety of specialists. Following discussions with typesetters, photographers, artwork technician and printers an estimate and a time schedule was devised. In a letter to the client the designer's understanding of the brief was presented. Also given was an estimate of the design and printing cost together with outline of the staged design activities, fee structure and schedule for delivery. The date for the next meeting was suggested and later confirmed by the client for the next week.

(vi) Synthesis:
After the briefing a variety of alternative compositions was intermittently considered. Over-night several solutions emerged and some were sufficiently resolved for later iconic renderings to be realised and appraised. Subsequently various representational possibilities were examined in terms of drawing style, illustrative technique and computer graphics rendering, etc.
(vii) Conceptualizing:

A suitable concept was required to reflect the overall message. Dunlop (1984) defines a concept as 'the singling out of general features that are easy to identify and can be used as the basis of experiments'. A concept is a single idea represented by a visual representation. It was essential to devise a concept which enshrined the client's message. The message was many faceted; a number of possibilities was devised based on attributes of convenience, simplicity, quality, user control, design potential, efficiency, cost savings, etc.

(viii) Creativity:

Creativity is in one respect the 'what if' stage in the design process. In another, it is the sometimes desperate search for originality. A number of preliminary ideas were informally discussed with colleagues. Later a formal meeting was convened with the client to explore the design possibilities more rigorously.

(ix) Brainstorming:

Brainstorming provides the opportunity to generate by free association, spontaneous responses often loosely connected in an attempt to devise an original and creative idea. Two designers who were otherwise uninvolved in the design task assisted in this activity. The notion was to generate as many design concepts and headlines as possible in approximately twenty minutes. There were no constraints or inhibitions; propositions could be humourous,
sensual, frivolous, contentious, conservative, formal, etc. Ideas could be sketched or verbal: the session was recorded for later contemplation. The correspondence was relaxed and twenty-eight ideas were evolved from which two recurring themes; 'label designing' and 'the finishing touch' was identified and eventually selected for development.

(x) Preparation of roughs:
By this time a variety of thumb-nail pencil and biro rough sketches had been prepared. These were examined and some were revised to conform with the devised concept. Roughs were drawn giving a variety of detail. A rough can give, at a small size and approximate proportion, an abstract impression of the position of the elements to be included, e.g. heading, image and text. The use of roughs enables the designer to rapidly and inexpensively recompose the layout to show and provide comparison of the proposed alternative configurations.

(xi) Evaluation:
There are no absolutes for the evaluation of the unique relationship between elements and codependents. The designer evaluated the various abstract propositions by consciously applying the agreed house style and production constraints, conventional rules of legibility, form, proportion and symmetry, personal preference and intuitive reasoning. As the design became more refined greater attention was applied to design details such as choice of typeface and combinatorial properties of alignment. A full size pencil rough
was prepared for discussion at the meeting with the client. All the elements were shown; image position, heading and blocks of type were represented.

With few revisions the design approach was agreed and the revised copy was presented. Additional narrative copy and specification was presented by the client for inclusion in the publication. It was agreed that this information should be accommodated on the reverse side of the publication. In terms of the design involvement, this was like a new commission requiring much of the previous consideration to be re-applied. A further layout was needed requiring all the constituent design considerations to be made as applied to the front of the publication. A revised production quote was required.

The client presented the requested sample of his brother’s artwork and it was agreed that the standard was appropriate for the desired line drawings. However the leaflet was now predominantly photographic rather than illustrative as the client had envisioned, so few drawings would be required.

It was several days before the various participants in the production could be consulted. In due course a further design, new costing and schedule to include the additional information was prepared. A new budget and revised delivery arrangements were subsequently agreed.

(xii) Commissioning photography for transparencies:

As this was a new product there were few opportunities to use in-house library
photographs so it was quickly realised that new images would need to be commissioned. The copy was examined to determine where photographs should be used to emphasise the concept and visually support the narrative. Sketches were prepared for the desired photographs. The photographs were commissioned and arrangements made for them to be taken. The product shots were taken at the company’s offices and a model for the female hand holding the label was arranged and the photograph was subsequently taken at the photographers studio.

(xiii) Illustrations:
The three line drawings required were prepared by the client’s brother using conventional line drawing techniques. Three interlocking circles of type were designed to show the design potential of the system. The type setting for this element was commissioned from a specialist computer graphics studio which used ‘Aldus Freehand’ to fashion the visual effect.

(xiv) Preparing copy for typesetting:
The copy preparation involved accurately counting the copy supplied to determine the number of characters to ensure that the type elements could be accommodated in the space available.

The type setter had previously been requested for and had provided specimens of the type style required. The type was marked-up to show type face, type size, leading, alignment specification and text alterations for the typesetting.
On return the galley (type set in single continuous column) of typesetting was checked and corrections were marked-up for resetting. The reset type was received and found to be without error. The sections were given an alphabetic code and their positions were indicated on the pencil rough.

(xv) *Transparency preparation:*

The transparencies are sized and marked-up using tracing paper overlays to show areas for proportional enlargement. Drawings made of the relevant areas of the transparencies were photocopied and retained for reference. The transparencies were coded for the printer's collection. On inspection it was determined that the transparency of the main label suspended from the hand should be retouched to reinforce the lines in the barcode; this was arranged by the printers who had been previously contracted.

(xvi) *Preparation of client presentation visual:*

At this stage a refined colour visual of the proposed publication is usually prepared. A client presentation visual as this is known, is an accurate colour rendering of the proposed graphic composition on which photographs, colour illustrations and line drawings are shown in the quality of the printed item. Title, heading and subheading are rendered to show the characteristics of the typeface. Text of a size less than 12 points is suggested in one of a variety of techniques representing the 'x' height of the type size, e.g. ‘Greeking’ or ‘tram lines’ etc. More refined visuals will use ‘body copy’ (text of the size and style intended but made up of nonsense word strings). A further
refinement is to use the actual live text setting if this is available. Increasingly colour visuals are prepared using computer graphics systems where available. However the production of a colour visual can be quite time consuming and therefore expensive.

The purpose of a highly finished colour visual would be to show the client a visual impression of the printed publicity leaflet. This would provide verification of the designer’s intentions before printing commences. However in this instance the client insisted that to reduce cost and save time a visual would not be necessary.

(xvii) Detailing: Preparing working layout.

The designer prepared an accurate drawing showing the position of the elements, heading, imagery, text and tables. This drawing on tracing paper is described as a ‘working layout’ of the design. Copies of the working layout were provided to the various contributors to ensure the same understanding of the prescription is achieved. It gives exact location of every element and can be used to verify the area accommodated by the elements. The various elements e.g. transparencies and text setting, are coded and cross referenced on the working layout. Areas of spot colour were identified.

The primary functions of a working layout are to aid the artwork technician’s positioning of the elements and to provide a check on the accuracy of the elements by the designer when the artwork is completed.
At the next meeting the designer presented the working layout together with the prepared transparencies and a photocopy of the text which received the client’s approval. A further meeting was arranged for the following week to agree the artwork before printing could commence.

(iiixx) **Artwork preparation**

The artwork was prepared by the designer from the prescription provided on the working layout. The artwork comprised the text elements, line illustrations, line drawings, box rules and spot colour. All the text and black working was positioned on the base board. Different colours were created on separate overlays. Areas of tone and spot colour were indicated on an attached tracing paper overlay.

The positions of photographic imagery were indicated in blue pencil which avoids reproduction difficulties while enabling ready identification. The artwork was carefully checked to ensure that all the information necessary for the printer was available. A cover was made to protect the artwork. At the next meeting the client examined the artwork together with the working layout which was approved after verification.

(iixx) **Production:**

Several sheets of colour proofs of the transparencies were supplied by the printer for the designer’s inspection. The printer had previously identified several colour variations; the designer inspected and marked-up a variety of
additional colour imperfections for correction. The collection of the colour proof sheets and the completed artwork by the printer, was arranged.

A set of revised colour proofs was returned by the printer two days later for checking. The checking of details included a visual examination of colour registration and a separate check for dust specks before approval by the designer.

The following week several sheets of page proofs were presented. The designer identified line breaks in two of the box rules which the printer agreed to correct. A meeting was arranged for the client’s approval of the printed page proof. The client requested a change in the colour of the ‘Olympic Transport’ logo. When advised of the cost of a colour change at this stage he desisted. The artwork was approved, stamped, signed and dated by the client and made ready for the printer’s collection.

(xx) Printing
A print run of 3,000 leaflets was ordered for delivery in 500 sheet parcels to the client’s office. On completion, six of the leaflets were delivered to the designer by the printer for inspection and filing. The designer met with the client for the final inspection of the leaflets before distribution. The result was approved; designer and client were satisfied with the success of the enterprise. However it was agreed that a presentation visual would have avoided the need for compromise at the proof stage. The duration of the design and production
process from briefing to printing had taken approximately three weeks. Much of the design time was spent in examining the design problem, generating ideas, evaluating and re-evaluating propositions. All costs were met by the client as agreed.

2.8.2 Glaze's model of the graphic design process

In this section the results of the graphic design case study are used to evolve a simple model of the graphic design process, see figure 2.5. Clearly the nuances of the universe of graphic activity are not represented in the elicitation of procedures used to produce an A4 size publicity leaflet. In this rudimentary example the procedure has been simplified, as follows:

1. **Client Briefing.**
   - Establishing the client's needs
   - Assimilation
   - Analysis
   - Synthesis

2. **Conceptualizing:**
   - Creativity
   - Brainstorming
   - Preparation of roughs
   - Evaluation

3. **Detailing:**
   - Preparing working layout
   - Preparation of client presentation
     - visual
   - Commissioning
     - Photography
     - Illustration
     - Typesetting
   - Evaluation

4. **Production:**
   - Artwork
   - Printing

*Evaluation occurs intuitively throughout the process. At the close of each stage a more rigorous evaluation procedure is undertaken; particularly at the detailing and production stages.*
Figure 2.5: Glaze's model of a traditional graphic design process.
Notable differences from the environmental design, architectural and engineering design models in this design procedure result from the idiom and idiosyncrasies of graphic designing.

2.8.3 Tunnicliffe's computer assisted graphic design procedures

Tunnicliffe (1991) reports a computer assisted procedure for graphic layout as consisting of the following activities, figure 2.6.

i. **Determine the image required.**
   - Assess the content.
   - Assess audience or readership.

ii. **Determine the appropriate size of the text to be used.**
   - Determine grid of graphic layout.
   - Choose point size and leading.
   - Assess readability.
   - Determine the amount of text.
   - Check to ensure that sufficient space is available for imagery.
   - Check the quality and quantity of imagery.

iii. **Choose style of text to agree with image of composition.**

iv. **Examine imagery.**
   - Check:
     - colour and tone information content relationship with image of composition
     - quality of imagery and their reproducibility
     - flexibility and balance
     - format
     - imagery - illustrations and photographs
     - tone and contrast
v. Determine rough graphic layout.
   Explore different configuration of the elements.
   Develop the most promising roughs.
   Evaluate the roughs devised.
   Refine the most promising rough.
   Add detail of title, heading and captions
   Evaluate roughs to ensure satisfaction with graphic layout.
   Rework if dissatisfied with design.

vi. Prepare full-size rendering.
   Check precision.
   Verify type size
     type style
     weight
     distribution and space
   Check alignment of elements.
   Accurately position text.
   Check text fits in grid.

vii. Filling in precise details.
    Location of imagery
    Treatment of captions
    Selection and placement of imagery
    Position of headings and sub-headings
    Consider use of rules; borders, dingbats etc
      use of drop initial
      use of colours
    Check exact placement of text
      adjust fit and position if necessary
      treatment of paragraph breaks
      location of paragraph breaks/widows
      indents and initial characters
      rules; tidy or improve the appearance
      of text and titling

viii. Detailed adjustments.
     Determine exact placement and identification
of photographs.
scale
cropping
fit with grid

Exact placement and specification of captions
Check use of rule to balance design
Make any aesthetic changes
gaps
positioning of sub-headings

ix. Marking up/type specification.
Specify text
Specify title
Specify sub-heading
Specify special typographic treatment
(colours, rules, indents, initial characters)
Identify photographs and their position
Mark position of elements e.g. folios

2.8.4 Cotton's computer assisted graphic design procedures
Cotton (1990) compares the similarities and differences in producing a conventional graphic layout and a 'Page Maker' graphic layout. He contends that the main benefit of the latter is an improvement in productivity due to the availability of type setting and layout facilities. His computer-aided graphic design model runs as follows., figure 2.7.

i. Initial briefing
Produce roughs
Choose type faces; interactive on screen selection.

ii. Input text to word-processor files.

iii. Proof text onto printer and correct.

iv. Commission imagery e.g., photographs and illustrations (these can be digitally produced).
v. Obtain prints and digital artwork:
Digitize into computer using a video image digitizer or flatbed scanner, scale, crop to size, retouch, manipulate colour balance, etc.

vi. Prepare graphic layout grid using a page-makeup program.

vii. Electronically paste-up complete pages:
Place typesetting.
Add tints, specifying Pantone or process colours if required.
Add required line rules, boxes, line artwork.
Check results on high-resolution monochrome or colour monitor.
Add full-colour images last (use monochrome versions of the colour image as a working guide for speed of processing).

viii. Output to ‘laser’ printer or colour ‘PostScript’ printer for hard copy, e.g. cannon.

ix. Make design corrections.
Prepare colour separations.
Output via imagesetter.
Identify transparencies to be electronically separated by repro house.

x. Prepare specification.
Inks, paper and order printed proofs from printer.

xi. Check printer’s proofs for imposition, colour balance and registration.

xii. Order print run.

xiii. Supervise required print finishing processes.
Varnishing.
Blocking.
Die stamping.

xiv. Supervise collating and binding, etc.

xv. Deliver to client.
Figure 2.6: Model of Tunnicliffe’s computer assisted graphic design process.
Figure 2.7: Model of Cotton’s manual and computer aided graphic design process.
2.9 Similarities and differences in traditional and computer assisted graphic design process

Comparison of Glaze’s, Cotton’s computer based and Tunnicliffe’s computer assisted models reveal consistency in many of their procedures. Indeed other than the difference in the number of steps represented, i.e. there are nine steps in the computer assisted process and fifteen in the computer based process, the main difference is that the computer based process facilitates the expeditious electronic configuration and production of a layout, while the computer assisted procedure requires the separate procedures of text specification and text setting to be completed as a prerequisite of artwork production. What is required is a system which not only facilitates text manipulation and artwork production but also provides routine design evaluation.

A variety of computer software is available to assist with the preparation of page graphic layouts. There are differences in the procedures and applications of these systems which affects their utility. Some desktop publishing packages are designed for the preparation of whole documents (document oriented) following a preset style, while others are devised to treat each page as a separate entity (page oriented). Page oriented packages require additional procedures to prepare or modify subsequent pages.

Some desktop publishing packages have the facility to position text in columns (column based) others provide a more flexible approach which ostensibly enables the designer to position text and image elements in predefined positions (grid based) graphic layout procedure.
A graphic design page layout may be based on a one column width or is more usually composed on a grid which subdivides the page into two or three vertical columns. A column based system positions text and image blocks routinely, starting from the top left to the bottom of successive columns working from left to right across the page. Any illustrations can be represented by a boxed area and manually positioned vertically within or horizontally across columns. The text can flow down the column avoiding reserved image boxes. In a grid oriented procedure the vertical columns are further divided horizontally to form rectangular boxes. Text can flow down the columns in the same way as in the column based system. The difference is that imagery can be positioned in the columns to align with any horizontal grid line. Imagery is represented as abstract boxed shapes just like those used in a column based graphic layout.

The productive use of computer graphics for desktop publishing relies on complementary expertise in graphic design and an appropriate computer program. Evidently graphic designers have by necessity embraced computerized procedures as an essential facility in their armoury of production techniques; respecting the need to maintain design standards equal to manually produced designs.

However, while the availability of computer graphics systems for desktop publishing offers the advantages of productivity and user control of the graphic production process, it is evident from the abundance of badly designed publications, that novices would be advantaged by the ability of desktop publishing systems to support routinely good graphic design practice.
CHAPTER 3: RULES IN GRAPHIC DESIGN

'Selecting a font absolutely in tune with the text; designing a consummate page with harmonically perfect margins, ideally legible, with immaculate word and letter spacing; choosing genuinely beautiful and graceful part-title pages in the same key as the text page - by these means a book designer can contribute much to the enjoyment of a valuable work of literature'.

see Tschichold (1991)

3.1 Introduction

Tschichold, arguably the most influential graphic designer, and foremost theorist (see Twyman 1992), wrote that 'typography is both art and science'. It is maintained in the literature reported that graphic designers who work with the strictures of type, book designers for example, abide by typographic traditions. Chapter 2 reported the interdependent nature of graphic design and typography. Accordingly, throughout this chapter the terms graphic designer and typographer are used synonymously.

This chapter discusses the rules and recommendations which graphic designers consider in the prescription of a graphic layout. It traces the graphic designer's traditional values, many based on centuries of good typographic practice.
According to Lauzzana and Pocock-Williams (1988) 'since antiquity, there have been numerous attempts to characterize artistic creation as a set of rules, witness the Pythagorean rules of the Golden Section. In the Renaissance, artists like Alberti and Durer formalized rules for projective geometry and ideal proportion'. The rules of the Golden Section remain the cornerstone to layout in architectural, engineering and graphic design, although not always explicitly prescribed.

Few graphic designers report their rules of graphic design and fewer of layout design specifically. Researchers such as Schenk (1991) and Tunnicliffe (1991) report that little is known about the graphic design domain. The work of Paul Klee (see Spiller 1969) is an enduring example of the Bauhaus teachers who attempted to analyse their work to provide rules and laws. But, some graphic designers will argue that there are no rules of graphic design e.g. Gill (1981). The paradox is that these authors also assert that graphic design is a creative discipline which relies on breaking of prevailing rules and conventions of graphic design!

Nevertheless it is evident that graphic designers have intuitively simplified their task and have made it more consistent by introducing written rules (in terms of house style and type specification, etc.) for the preparation of certain repetitious procedures. The observable uniformity of the configuration of the layout of the pages of books and magazine pages, for example, show that the structure of the layout and the size and situation of its components have been
prescribed for routine distribution. Knuth (1984) reports the \LaTeX\ system in which is embedded certain rules of the logical structure of typographic design, see also Lamport (1986).

A survey of the literature revealed that while some graphic design prescriptions are explicitly stated many are implicit and intuitive. Intuitive prescriptions are manifest as recommendations rather than explicit rules.

The main source of graphic design recommendations relevant to this domain proved to be the conventions of book publications. The information that follows is based on accounts given by book designers and researchers into book design and typography. Influences on the graphic designer's choices and the characteristics associated with the physical properties, e.g. symmetry and asymmetry, organisation, structure and legibility attributes of a graphic layout in this domain are examined.

Where graphic designers have given a comprehensive account of their recommendations, these have been quoted. The views of some graphic designers and researchers are intermittently referred to in support of particular rules to highlight their agreement or disagreement.

3.2 Factors of typographic composition
There are several factors which affect a typographic composition. Normally a graphic layout conveys both information and style through the choice of
elements and their configuration. The nature of the information in a graphic layout can be very variable. The graphic designer chooses the appropriate founts, for example, to match their emotive response to the information. Assuming the availability of the selected typeface, the type size is conditioned by the physical constraints of paper size, format, columns and margins.

The elements in a graphic layout are arranged to enable the reader to routinely access the information contained. This is achieved through the establishment of priority, emphasis and sequence of the content. The graphic designer provides a clear path from title to footnote ensuring that priority is well signposted in the use of headings and sub-headings throughout. Additionally the choice of type for title and text and the size and weight of type are the discretion of the experienced designer.

The style (structure, configuration and choice of type elements) of a layout is the consequence of the graphic designer's interpretation of the information. Style may reflect the graphic designer's attitude to the content. Style is observed in the configuration and structure of the layout; symmetric or asymmetric. Style in a layout may be formal or informal, classical or contemporary and can be expressive.

It is evident that a graphic designer may seek to overlay their personal preferences onto the information. The graphic designer's preferences are observable in the layouts of experienced individuals. A graphic designer may
produce a layout which underlines rather than overlays, the narrative information of a page. Tschichold (1958) advised that 'the more significant the book, the less room there is for the graphic artist to position himself and document through his 'style' that he, and no one else, designed the book.'

Zachrisson (1965) reports that typographic compositions have two conflicting distinctions; the classical and the expressive style. The first, he suggests, is typified by the work of Stanley Morrison who is reported to maintain that 'it is better for typography to be uninteresting and monotonous than to risk distraction in the act of reading.' The second may be described as Bauhaus typography which has the distinction of being asymmetric in composition, sans-serif in preferred type elements.

3.2.1 Distribution of elements in a page layout
In the first instance the proportion of the page is usually determined by commercial or production constraints. Secondly, the contents (imagery and text) are usually provided with cursory regard for their distribution. Nevertheless, the distribution of elements within the page demands particular attention to ensure its balance within the confined space. Rubinstein (1988) suggests that the shape of the content should match the shape of the page, e.g. if the orientation of the page is wide (landscape) then the rectangle within which the content is described should also be wide. There is evident variation in the density of elements in a graphic layout, Tufte (1990). This is observed to constrain the distribution of the content. The ratio of space to content is
inhibited by both format and size. Inadequate space inhibits the orientation of elements. While limited space restricts size of text which affects legibility and readability, the availability of space permits the optimum choice of type size and effective distribution of elements. According to Tufte abundant space provides complementary configurational challenges.

A one-page graphic layout may consist of a headline, illustration and copy, where the headline is the 'attention-grabber', the illustration reinforces the story and the copy imparts the message, Shepherd (1987). Additionally headings and sub-headings may be included in a graphic layout enabling the reader to quickly identify specific information, e.g. venue, date, price, speaker, etc.

Furthermore special founts may be used to reinforce a particular quality, merit or attribute. The attribute of each element is assigned to fulfil a predetermined need. Characteristics, such as the type size, leading, number of columns, column width and margins contribute to the accessibility of the content.

3.2.2 Formal structures and grids

Structures

Many artists are concerned with establishing connections between natural forms and artificial structures; 'The notebooks of Leonardo Da Vinci' (1452 to 1519) edited by Richter (1970) being a comprehensive example of an artist's researches in this regard. Mathematicians have also thought there was a
relationship between natural forms and human structures and sought to
represent their beliefs numerically. Leonardo Fibonacci a notable
mathematician, devised ratios such as the Fibonacci series 1/1, 2/1, 3/2, 5/3,
8/5, 13/8 etc. to represent the spiral of leaves on the stalk of a plant and to
provide a better understanding of natural forms which might inform human
endeavour.

Arguably one of the most important ways of classifying a design is by its
structure. For example a design may be described as being symmetric or
asymmetric.

Structure in fine art

Klee (see Spiller 1969) came to believe that there was a natural hidden
numerical relationship of parts corresponding with the whole in both man made
works and in nature. He developed a theory of pictorial form and studied the
receptive process, gravity, colour and dynamics in visual imagery. According
to Klee the creator’s first action dictates the centres of energy in a composition.
He studied the way the eye moves across an image and concluded that the eye
moves from strength to contrast to settle on the weaker values.

Painters like Mondrian experimented with the division of a canvas into ideal
proportions of both colour and space. Gombrich (1965) speculates that these
artists influenced op and pop art. The threads of abstraction are evident in
many forms of graphic communication, contemporary publications, computer
displays and the printed page.
Structure in typography

In the 1920's work on structure, form and function at the Bauhaus school in Germany greatly influenced aesthetic attitudes in architecture and other design domains. The work of many individuals associated with the Bauhaus like Paul Klee (see Spiller 1969) provided the foundation of a contemporary visual language of graphic design which became established principles and norms.

Their initiatives in experimental typography evolved a departure from the prevailing preference for serif typefaces and began a proliferation of sans serif faces for posters and notices which challenged traditional practices. This resulted in a move away from the traditionally held belief that text books should be set in serif faces.

Against the 'rule' that text should be set in upper and lowercase they experimented with text set in lowercase throughout. Tschichold (1991) became an ardent disciple of the Bauhaus movement. Initially he advocated what acquired the name of 'modern typography' and 'asymmetric dynamism' and suggested replacing all the German types with a single sans serif style. He abandoned the norms of symmetry for asymmetry and expressed his radical convictions in typographic articles, his book 'Die Neue Typographie' and his numerous experimental posters. However, he later acknowledged that a variety of typographic styles was required to serve the needs of a mass market. His design prescriptions gained in popularity and through the subsequent decades his principles and those of his peers have endured and their designs
imitated; graphic designers have adopted and adapted many of their values, inherited many of their prejudices and endorsed their preferences.

**Structural anarchy**

According to Burden, Morrison and Twyford (1988) typographic anarchy is epitomised in the words of futurist Filippo Tommaso Marinetti, who wrote ‘Our revolution is directed against the so-called typographic harmony of the page, which is opposed to the flux and reflux, the jerks and bursts of style that are represented in it. We shall use, therefore, on the same page three or four different colours of ink and, if necessary, even twenty different forms of type’.

According to Reed (1988) in the 1920's Dada emerged as a reaction to the setting of prescriptions. There was a strong move to break with prevailing traditions. Notable of their beliefs were that firstly ‘we should rebel against the tyranny of the words harmony and good taste’ and secondly that ‘art criticisms are either useless or detrimental’.

In essence Dadaists proclaimed the abandonment of rules. But it is unclear what they hoped to achieve; perhaps popularisation and acceptance of their beliefs in artistic anarchy. Their aspirations had little to do with quality, legibility and readability but more to do with change for its own sake and the creation of an acceptance that anything and everything could be called art. According to Wosencroft (1988) ‘in a certain sense, art died with Dadaism - with Duchamp and his “ready mades”.’ Today the renaissance of Dada can

63
be detected in the early work of such contemporary graphic designers as Neville Brody (see Wosencroft 1988).

Personal preferences and expertise predispose us to a particular range of design options, solution types. Styles change and with them designers, client and audience expectations; though not necessarily simultaneously.

The availability of new technology like computer graphics and desktop publishing systems like Quark Xpress and PageMaker provides the opportunity for new approaches and offers the potential for further experimentation in the search for new solutions and the departure from earlier stereotypes and convention. However, these systems rely on the use of a grid system to ensure structural integrity in the composition of an individual piece and serial uniformity where a document consisting of two or more pages is being prepared.

**Grid systems**

Typographers use the grid system to subdivide and configure their graphic layouts. Meggs (1983) reports that there is evidence of formal structures such as the grid system in the early Sumerian pictographic tablets, c.3100 BC, in which ‘information is structured by horizontal and vertical division into zones’. Rotzler (1988) argues that people have always been preoccupied with systems and structures to support their visual expression. In support of this proposition he reports that ‘a so called geometric-style’ is evident in all the
'great early cultures.' He reports that it is 'exemplified particularly in New Stone and Bronze Age pottery of the fifth to the second millennium BC in Egypt, Mesopotamia and above all Persia'. Researchers associate this with the evolution of exact measurement of time and what mathematicians call 'mathematical ornament'. Many authors contend that this is evidence of our early preoccupation with formal structures.

Notable among these systems is the Golden Section which has been used extensively in architecture and publishing. The Golden mean is an expression for the classical ratio (approximately 8:13) created when a line is divided into two parts in such a way that the ratio of the shorter to the longer is as the longer to the whole. Its use can be observed in the external proportions of ancient buildings like the Parthenon built 432-447 BC and provided the ideal proportions for the image area of a printed page of antique publications like those produced by Gutenberg and others in the 15th century. The economic constraints of the present day inhibit the use of such proportions.

However there is confusion in the use of the terms Golden mean, Golden rectangle, Golden Rule and Golden Section in some domains; figures 3.1, 3.2, 3.3 and 3.4 illustrate my usage.

The Bauhaus brought the era of structuralism in which designing became increasingly focused on form and function, and in which two and three dimensional simplifications were pursued in search of 'ergonomic' suitability.
Similarly, many artists have used a grid to scale sketches and cartoons to fit the proportion of larger works; grids are essential to mural art. For centuries, plans for navigation and military action have been plotted on latitude and longitude grid co-ordinates by cartographers. Classical architects used grids to prepare scale plans and elevations for the effective division of space plans for human convenience and appearance.
Hurlburt (1970) observes that ‘wherever plans have been called for in the building of objects, the division of areas, or the decoration of flat surfaces, grids have been involved.’ It is evident that the scribes since antiquity used grids to ensure the conformity of margins and the ‘house style’ of their manuscripts was maintained. Since Gutenberg, typographers have used grids to organise graphic layout elements on the printed page.

A typographic grid divides the page area into vertical columns and horizontal rows. The column width can be determined by the application of known prescriptions which give 8 to 12 words per column for optimum readability, Hurlburt (1977). The arrangement of the horizontal subdivisions regulate the distribution of the text and image elements across the page. In general the grid is composed of a series of even or unevenly spaced parallel horizontal and vertical lines for the location of elements to be incorporated in the graphic layout. The outer vertical lines control the front and back margins while the top and bottom horizontal lines control the head and foot margins of the page.

A grid shows the vertical and horizontal margins together with internal cells.

Conventionally the side margins on a single page are equal while the top margin is smaller than the margin at the foot of the page. According to Tschichold (1991) the ideal ratio of top to bottom margin (for books) is 1:3; side margins should be equal to the top margin.
Figure 3.3: The Golden Rule $1 : 1.618$ is applied in this example to the ideal proportions of a page; it does not give the ideal position of text areas according to Tschichold. The text area is proportional to the Golden section.
Figure 3.4: The term Golden Section when applied to the pages of a book, relates to the division of the format to form ideal text areas in Tschichold’s opinion.

Twyman (1992) postulates that the conventional proportions of book margins are rooted in the bilateral symmetry evident in hand written sixteenth century lithographs. He attributes the ‘conventional’ ratio of ‘side margins’ back and fore edge (left and right edge respectively) to functional rather than aesthetic exigency. The subsequent developments in printing and the availability of new technologies, he suggests, facilitated the challenge to established graphic layout styles. Subsequent developments were to have significant effect. The
introduction of typewritten copy for example, promoted the gradual transition in page graphic layout from bilateral symmetry to asymmetry.

Garland (1969) gives the ‘constraints on the placing of type area on loose-leaf sheets’ and ‘some alternative column arrangements’. He graphically illustrates and specifies his criteria for deciding on the subdivision of type areas into columns to accommodate tabular matter and side headings, and gives recommendations for the relative height and width of head, foot, back and front margins, figure 3.5.

Hurlburt (1977) compares a variety of grid-structures contrasting their flexibility and application. He reports the ‘orthodox design grid’ which became popular in Europe after the second world war which is based on the division of the graphic layout into uniform squares. This allows a natural pattern of squares and the combination of pairs of squares to form rectangular blocks. Other examples include one designed by Alan Fletcher for 24 Ore which features ‘strong horizontal emphasis that is characteristic of many newspapers.’ He describes another devised for a Mitchell Beazley encyclopaedia by Ed Day to ‘cope with the widely varied content in such a volume.’ see figure 3.6.

The structure of a graphic layout may not be apparent to the casual observer. The configuration of elements within a graphic layout may be disguised by the content or emphasised by its strong horizontal or vertical structure.
Figure 3.5. Garlands type constraints on an A4 page (Garland 1969)

Page constraints on the placing of type areas on loose-leaf sheets

First constraints: back (binding) margin

Second constraints: head and foot margins

Third constraints: fore-edge margin

Fourth constraints: heading margin

Two equal columns

Three equal columns

Two unequal columns

Four equal columns

Loose leaves may be bound by plastic clips, plastic ring, post or ring binder and by multiple ring binder; in each case a margin of not less than 0.63 should be allowed.

The minimum margin at the head and foot of the sheet should take into account inherent tolerances: a recommended minimum is 1/4" (1/2", max).

As with the head and foot margins, a trimming tolerance should be allowed for, and here also 1/16" minimum is recommended.

Where headings are frequent, as with catalogues, it is as well to allow a heading margin in addition to the head margin, even though it may not always be occupied by a heading; the size of this margin will vary according to the kind of heading and the size of page.

The most common arrangement, giving a reasonable line length and a simple layout structure.

Allowing greater flexibility than the two-column grid, three columns may be used either for continuous reading matter or for tabular matter, and small illustrations can fit conveniently into the grid.

Two unequal columns may be useful where side headings and hanging notes are required; the grid above is suited for using part of the structure of the three-column grid: this gives more flexibility than a grid of two unequal columns which does not relate to an equal column grid.

A four-column grid with tabular layout and column function on a two-column grid or the left-hand column can be used in conjunction with a wide column the same width as the other three columns taken together; it may be advisable to compose a two/four column grid of this kind even when only two column use is at first envisaged.

General note: every form of publication has constraints on the placing of type and illustration, which result from its printing, its binding method or the way in which it is intended to be read, filed or retrieved. The example above is only one of many such situations, although perhaps it is the most common.
Figure 3.6: Ed Day’s grid designed for a Mitchell Beazley encyclopedia
Swann (1969) and Ballinger (1970) support the soundness of devising a grid within which a formal graphic layout may be composed. They concur that use of a grid provides unity within a publication and harmonises the proportions of a series of graphic layouts based on the same format. A designer may employ a classical grid 'through the application of certain systematic principles of proportions developed by mathematicians, artist-designers, and architects throughout the course of design history' reports Hurlburt (1977). Swann C (1969) provided a comparison of the classical principles of the 'Golden Section', the 'Golden Rectangle', the 'Golden Mean'. Swann A (1991) describes the contemporary approach to dividing the page. By these methods it is possible to determine the size of the main type area, columns and margins on a page of any given shape.

Graphic designers are divided about the suitability of using such formal tools. Many designers, like the architect Le Corbusier who developed a proportional system he called the Modulor, advocate the use of formal systems, (Le Corbusier 1960). However they reserve the freedom to depend on feelings rather than system and reason to answer design problems. It is also acknowledged that design grids may be a recipe for the mundane when used by the inexperienced. Others describe the grid as a strait-jacket useful only to magazine and book designers, but there are many of these! (Hurlburt, 1977.)

Advocates of the grid system, like Hurlburt (1977), affirm that it is particularly appropriate for the design of a series of graphic layouts, to a common theme or
soundness of devising a grid imposed. They concur that use harmonises the proportions of nat. A designer may employ remain systematic principles of sist-designers, and architects Hurlburt (1977). Swann C il principles of the ‘Golden Mean’. Swann A (1991) the page. By these methods pe area, columns and margins of using such formal tools. who developed a proportional use of formal systems, (Le cedom to depend on feelings ign problems. It is also r the mundane when used by a strait-jacket useful only to of these! (Hurlburt, 1977.) ), affirm that it is particularly youts, to a common theme or ne, journal or book, even sign systems and there is unanimity that a virtue of the grid ses on the untrained designer. He the graphic layout may result in a more an experienced designer.

specialises in advertising graphic layouts, coach for the composition of such layouts. warns that a ‘grid system is an aid not a r (1981) ‘proportional systems based on work.’ Borgman, Muller-Brockman, ith one-off compositions a grid may be means ‘well-proportioned, well-balanced, oncordance of several parts by which they bound up with symmetry’. He quotes results from proportion ... Proportion is is constituent parts with the whole. He mean correctness and is preferred by many ern examples are given to underline the architecture, sculpture and painting.
The nuances and attributes of typographic symmetry have been considered by Tschichold (1991), Twyman (1992), Goldring (1966) and McLean (1975). It is Tschichold’s view that ‘there is no such thing as symmetrical typography’. He explains that ‘the word symmetry may not be used when we talk about typographical arrangement, because something is symmetrical only if one half is the mirror image of the other half.’ Symmetry is by his definition unattainable except in centred imagery or centred text. The term ‘centred typography’ he suggests more accurately describes a series of lines which are set with equal space at opposite ends.

Tschichold (1991) suggests that ‘good typography has a simple structure. The centred line is a specific and indeed supremely important structural component in good typography’. He contends that ‘to centre lines of different weight and type size one beneath the other is at the same time the simplest and the best typographical method’. But centred lines of type do not produce ‘symmetrical typography’. Symmetry is attainable only when applied to centred imagery.

Twyman (1992) reports the use of bilateral symmetry in the design of the double page spread. He describes bilateral symmetry as the mirroring of two facing pages of a book, and gives reasons for its popularity from medieval times. He contends that ‘the argument for symmetry in book design is more convincing than any that can be applied to symmetry in, for example, architecture’. Asymmetric design he reports became accepted in the twentieth century by British designers, typified in the work of Antony Froshaug and
Herbert Spencer. He suggests that the design of book pages results pragmatically from the 'convergence of technologies and domains'. He identifies the electronic typewriter which facilitates type setting by typists rather than compositors as contributing to the change to asymmetry in the organization of book pages.

Goldring (1966) examines the traditional preference for justified typesetting and centred headings in periodicals and books, etc., which results in symmetry in the appearance of the layouts produced. He reports that bilateral symmetry is favoured in many printing firms where it is believed that 'only bilateral symmetry can be aesthetically pleasing'. However, he emphasises the impact of 'non-justification as a functional requirement' of computer typesetting. Computer composition routinely expedites asymmetric, unjustified (left justified) text setting; justified type setting demands to be specified (changed from the standard unjustified by selection from various alternative alignments e.g. centred, left justified, justified and right justified). The practice of unjustified typesetting he observed, could predominate not least because it is practical; it requires fewer key strokes, and is therefore more economically produced than justified type.

3.3 Readability

Zachrisson (1965) asserts that 'text must be read, whether the publication is mainly conceived as a beautiful show-piece, or a modest specimen'.
Words form picture blocks which are uniquely recognizable by the reader. The characteristic of letterforms and their spacing may enhance or diminish the legibility and hence the readability of the word, or string of words. Research shows that the design of letterforms contributes to their recognition, Rubinstein (1988). Tufte (1990) reports that complexity enhances the identification of word pictures; ‘the more letters are differentiated from each other, the easier is the reading.’ The uniformity of sans serif letters reduces discrimination and produces poor word pictures.

Biggs (1949) distinguishes ‘readability, as meaning that an inviting, pleasant aesthetic is added to the mere recognizability. In this sense it would be possible to describe a type as being legible but not readable.’ For example we may be able to recognize the individual characters in a word string yet if the spacing between words horizontally is greater than the line spacing the observer may experience difficulty in reading the text.

He points out that ‘type has to be read with as little discomfort on the part of the reader as possible. But reading at different times is carried on under different circumstances, and the qualities desired on one occasion might be irrelevant on another. For example, in an advertisement, one or two words might reasonably and appropriately be in a bold or decorative or flourished letter, in order to attract attention. But the very qualities which are successful on such an occasion would be quite out of place in a book where continuous reading is the intention.’
Cues which signal changes in information

In a graphic layout, different cues can be used to inform the reader of a change in the flow of information. Where imagery is used, this may provide the primary focus, demanding the observers attention. The position of the title or heading signals the starting point for the reader. A change in typeface, type style and or in the weight of type may be used to differentiate heading, sub-heading and text copy. An increase in the size, and type style may be used to emphasise the importance of the heading but is undesirable for a sub-heading where a change in weight is more commonly employed.

3.3.1 Psychology of perception

The psychology of perception is important to graphic designers because they can exploit the effects. Albers (1971) reports that ‘we do not read letters but words, words as a whole, as a word picture. This was discovered in psychology, particularly in Gestalt psychology.’ The Gestalt laws of organization were formulated by Wertheimer. According to McKim (1980), the term ‘Gestalt is a German word that has no exact equivalent in English’. But it is widely accepted as meaning whole.

Our sense of order results from the growing experience of our environment and from the structure and function of our visual apparatus, according to Gregory (1967). Gestalt experiments, he relates, reveal that we, perhaps instinctively, tend to group visual elements to form patterns, rhythms, order and series. This natural facility can lead to a degree of certainty which Zakia (1975) says
enhances aesthetic experience. He reports that 'Gestalt laws of perceptual organization' relate to the conditions which regulate the stress and stability in a composition such as regularity, uniformity, simplicity, closure and symmetry. We are also reminded of the need for variation and that too much order or repetition (certainty) in an image results in the mundane: complexity without order produces confusion; order without complexity produces boredom.

Gregory (1967) reports that the brain is continually seeking to organize sensory data into objects. Often we see objects which do not exist; 'faces-in-the-fire' for example. The graphic designer seeks to exploit this facility by grouping or separating objects, enlarging or reducing to create emphasis, and alternating groupings to create rhythms and patterns.

3.3.2 Legibility

Tufte (1990) argues that the primary criterion in typography is to ensure the legibility of the information represented. Information which is illegible is not decodable, i.e. readable.

Rubinstein (1988) suggests that 'legibility is the degree to which text is easy to read. The qualities of text varies widely, from barely decodable, to text that is read quickly and comfortably by all readers'. So that for text to be readable it has to be legible. McLean (1980) describes legibility as a 'dangerous - and interesting - word. It is dangerous because it is often used as if it had a definitive or absolute meaning, which it does not have.' He contends that 'if
we say something is legible, we mean that in our opinion or experience the
people we want to read it can read it in the conditions in which we think they
will see it.' This suggests that text may or may not be legible to section(s) of
the general population. It is evident that many elderly people find it difficult
to read text which is in a small type size. So legibility is both subjective and
is moreover variable over time for an individual. What is legible today may
not be legible tomorrow. McLean (1980) gives three essentials to achieve
legibility:

i. Sans-serif type is intrinsically less legible than serifed type.

ii. Well designed roman upper and lower-case type is easier to read than
any of its variants, e.g. italic, bold, caps, expanded or condensed.

iii. Words should be set close to each other (about as far apart as the width
of the letter 'i') and there should be more space between the lines than
between the words.

According to Goldring (1966) the professional typographer is traditionally
cerned with appearance and legibility: appearance observed in terms of
recognizability of style and legibility measured in terms of a fount of
alphanumeric characters recognizability. More often it is measured in terms of
the recognition of strings of characters formed into words.

Spencer (1968) states that 'legibility in printing is concerned with the efficiency
of the visible word.' He reports the work of many researchers on the theme
of type legibility with particular regard to type size, type style, line length and
leading. He examines the arguments against type justification and the use of

80
capital letters in text matter. He concludes that type justification diminishes legibility and three or more lines of capital letters diminished readability.

Vertical and oblique 'text which departs from the normal horizontal alignment' are also examined by Spencer. He reports that titles arranged from head to foot were no more legible than those which ran upwards. However it is evident that words formed of strings of characters vertically aligned are relatively difficult to read.

It is evident that legibility contributes to and is conditioned by many factors some of which are reported above; some other codependents are reported in sub-section 3.3.3.

3.3.3 Column width and optimum line length

The width of the vertical columns determine the number of words per line in the chosen type size. Conversely, the number of letters per line can be used to calculate the range of permitted type sizes (see section 3.3 on readability).

Ruder (1981) reports that the number of characters in a line will influence the eventual legibility of a graphic layout. For example, while a line containing between 50 and 60 letters will be easy to read, 'too great a width of composition degenerates into a decorative grey and is not conducive to reading.' Furthermore he asserts that readability is enhanced by the reader's ability in picking up the beginning of each subsequent line. Where lines are
too short then reading becomes wearisome. This is supported by Spencer (1968) who reports the work of Javal, a researcher who in the nineteenth century determined that an increase in regressions (the inability to accurately determine the next line in a block of text resulting in the necessity to refer back to the end of the previous line) and pauses (wide word spaces resulting from narrow columns and justified type setting) slow down readability.

Another feature of a narrow column or short line length is the uneven spacing between words which results and may affect the grouping of words vertically to form 'rivers', a view supported by Shepherd (1987). Rivers present as continuous vertical empty spaces in a block of text which is sometimes greater than the width of adjacent word blocks. It results from the coincidence of spaces between words on a series of vertically adjacent lines in a block of text.

Twyman (1970) reports that reading methods have a certain constancy which has meant that although line length 'may vary from one book to another according to its size and purpose' the number of words on a line has remained consistent over the centuries. Nine words per line is thought to be the average. However this is an approximate figure which is conditioned by the use of technical terms and colloquial language in the text. Ruder (1981) like Spencer (1968) supports the view that 'line lengths can be varied within broad limits without diminishing readability but research has shown that very short lines slow perception'.
There is observable agreement that the optimum line length is between 50 to 70 characters. The tolerance is attributed to variations in type size, anomalies in the ratio of ‘x’ height to body size, and associated with the serif and sans serif characters, ascender and descender of different founts. The minimum line length is thought to be 40 character or six words per line, Swann C (1969). Hurlburt (1978) provides the prescription of 15 pica line length, 9 point type size giving 45 characters per line.

3.3.4 Variety in type faces and type sizes
Tschichold (1991) contends that ‘no author should use more than four different fonts in any text.’ Today there are estimated to be some 10,000 type faces available to the typographer. They range across five increasingly indistinct families and are designed for display and text copy. Each is designed to satisfy a particular need and evoke a discrete ‘tone of voice.’ It is self evident that there is no ubiquitous type face or type size. Consequently the problem for the novice is which fonts to choose.

Typefaces may be divided into two groups: the serif and the sans serif. Serif typefaces are derived from chiselled Roman letterforms. Serif typefaces are variable in thickness of stroke. A serif typeface is characterized by a horizontal feature at the base of the vertical features from which the letterforms are drawn. Sans serif typefaces are uniform in stroke and thickness. They have a mechanical appearance and are appropriate to contemporary subjects; reported in Tschichold (1991).
Muller-Brockmann (1971) reports Tschichold's (1925) pronouncement that 'typography has no use for a diversity of characters, a superfluity of type sizes, spaced lower-case letters, capitals for texts, ornaments and variously angled lines... Like Bayer at the Bauhaus, Tschichold championed sans serif, which has been popular with modern designers since 1925 because of its clarity.'

However, Albers (1971) contends that the fashionable preference for sans-serif in text shows neither historical nor practical competence. First, sans-serifs were designed as letters not for text but for captions, when pictorial reproductions were introduced with stone lithography. Second, they produce poor 'word pictures.'

Type is usually available in sizes varying from 6 points to 72 points. The size normally used is 12 points but this depends on typeface legibility and application. Rubinstein (1988) recommends the use of three to four type sizes for normal book pages. Setting in one size only is far worse than setting in too many fonts in one of Tschichold's rules, Tschichold (1991).

3.3.5 Capital letters
Albers (1971) contends that 'words consisting only of capital letters present the most difficult reading - because of their equal height, equal volume, and, with most, their equal width.' According to Albers text should not be set in all caps because their uniformity reduces readability or the speed of reading (see readability). This is informed by anecdotal reports typified by Alan Haley.
who in ITC's *U & lc* (Upper and Lowercase) journal (Fall 1991) reports that the
obvious reason for choosing upper and lower case letters is readers are more
familiar with them and research shows that they are more comfortable with
them. Haley reports studies which show that the variation in the height of
lowercase letters by their ascenders and descenders are an aid to character
recognition. The distinctive word shape or 'word picture' can be recognised
more easily and consequently recalled much faster than type set in all caps.

3.3.6 Type setting

Another contentious area concerns the setting of type. While Ruder (1981)
acknowledges that 'ragged edge composition' results in even spacing and a
uniform grey, he contends that *justified setting is preferred to ranged left* (left
justified) type setting because it is more familiar to the reader. Moreover he
argues 'where there is a great deal of text, reading is slowed down' where type
is unjustified.

However, Spencer (1966) argues that justified type setting results in a variation
in word spacing from one line to the next and the frequent breaking of words.
Carter, Day and Meggs (1985) assert that 'There are appropriate reasons for
setting justified or unjustified typography, but type set flush left and *ragged
right promotes greater legibility'. Shepherd's (1987) compromise is described
as semi-justification; 'it requires *all lines* that fall beyond a pre-determined
point within the measure to *be justified*, whereas those that fall naturally short
of the specified point *remain unjustified*.'
There is broad agreement that italic should be reserved to emphasise key words within a sentence; book titles etc. While bold should be confined to headings, and medium in some fonts used for sub headings. Text, usually the bulk of the copy, should be set in light or medium.

Type may be set with additional space between successive lines in a block of text. This additional space between lines of text is referred to as leading. Where type is set solid i.e. without any additional space between lines in a block of text this is referred to as 'set solid.' Where a type size of 10 point is used, additional spacing of 2 points, for example, may be added to improve readability. While too much or too little space between words can seriously affect legibility, Shepherd (1987), where a group of lines are closely spaced vertically, readability is also affected.

3.4 Production constraints

The production process is a major constraint on the format and consequent organization of the content, its component legibility and the resulting readability. Here, Garland (1969) suggests that the method of production and use imposes particular constraints on the design of a publication as it does with every form of design. As an example of this he reports that wide inner margins are required for comb and plastic binding while saddle stitching permits narrow margins to be employed. Other production constraints also affect the potential for variability in a graphic layout. Some printing equipment cannot process areas of solid black while on others there is a
restriction on 'bleed'. Many printing machines require a margin of at least 3mm on left and right edges.

Similarly the method of display, storage and retrieval also constrain the distribution of elements and the proportion of the format area in the composition. A leaflet is required to be read either on a notice board or held in the hand. It has a dual role; to be seen at a distance as a small poster, while possessing the characteristics of a book page which is usually viewed at close proximity.

3.5 Can we define 'good graphic design’?

Aesthetic evaluation in the graphic design domain is concerned with determining the suitability of the encoded representation. The aesthetic evaluation of a graphic layout has two components; emotive and physical. Emotive evaluation is concerned with the affect produced in the recipient of the visual representation. Physical evaluation is concerned with the individual's judgement of the association of the elements in a composition. It is also concerned with the mental comparison of the visual representation with the recipient's expectation based on their preferences in terms of structure and organisation for example, and experience of similar solutions in the population.

A primary concern in this thesis is the notion of good design and whether a rigorous definition can be formulated for the aesthetic. According to Lewis
and Brinkley, (1954) 'we can say that a successful piece of graphic design must be logically conceived, must be fashioned with taste and must be skillfully executed'. They maintain that 'it is impossible to give a more precise guide to what is or is not good design'. However, design may be judged against proven standards and established conventions. The use of a certain range of colours in fashion, textiles and furnishing in a particular decade or year gives a clue to the style and colour values of a generation. Good design can be judged against the properties in an available item. The qualities of the classical and enduring can be assessed and used in the evaluation of contemporary designs. A list of particular attributes can be assessed and used to judge similar features in other designs, old or new. The form, surface texture, structure, proportion of a piece of design are witnesses to the designer's preferences. Nevertheless, design evaluation based on rules formulated retrospectively will be incomplete because the prescription from which it was created is unavailable, and therefore cannot be used to comprehensively evaluate the universe of prospective works.

There is an observable commonality in the properties addressed in the literature where attention is primarily focussed on book design, see Tschichold (1991) and Rubinstein (1988). Here good legibility and readability and their interdependence on margins, type size and setting, and variations in type faces and number of type sizes, etc. have achieved a level of consistency. Consequently a degree of consensus can be extrapolated from these sources about the elementary rules of good design for book publishing.
According to Benton, Benton and Sharp (1975) good design must fulfil criteria relating both to its aesthetics and utility. Good graphic design equates to suitability of visual appearance, and utility to the ease of access to the information, e.g. readability and sequence. If a graphic design solution for example is visually stimulating but illegible it will fail to communicate its message to the reader. On the other hand if the content is legible but unattractive the reader might fail to notice it. The suitability of a graphic design solution resides in the balance between legibility and visual appeal.

Biggs (1949) suggests that 'Once you are agreed as to what a thing is for, the goodness or success of the thing will depend on whether or not it fulfils the purpose for which it was intended.' If the purpose of a traffic sign is to warn motorists of a local hazard then its goodness can be determined by whether drivers encountering it are able to decode the content in time to act on the message delivered.

However, in designing a video label or magazine cover the graphic designer may be more concerned about novelty than aesthetics and utility. Often style will influence the designer’s choice of imagery and the balance created between impact (visual stimulus) and information (content).

Individual designers have different values; what is ‘good’ for one designer may be considered by another to be ‘bad’. Therefore the notion of a universally accepted ‘good’ graphic design solution, relies on all the recipients of a visual
communication having the same perception of a designer's personal prescription and purpose. It is improbable that an automatic design facility could unequivocally advise of good design in the absence of a consensus on why a design is good. Furthermore, the permutation of designers' values currently cannot be captured comprehensively. Moreover Pye (1988) warns that 'To say things which have the following characteristics are beautiful is always an untruth' perhaps because they may also contain undesirable, ugly, characteristics.

A design cannot be judged good by a checklist of features because the list can have negative features:
• a good design has all of these...
• and none of these features...

Even if a list of features were sufficient, there appears to be no consensus on what it should contain. The problem is that even if a design has only desirable individual features, the interaction of these features may lead to 'bad' design.

3.6 Can we define 'bad graphic design'?
Rather than focussing on the merits of a design solution the alternative approach is to address the symptoms which cause 'bad' design.

A bad design can be defined as one containing a defect in either its aesthetics or utility. For example, the serif type used in a layout may be legible (which is not bad) but its singular use may result in 'bad' design. Also, the presence of
only one font in a layout can be specified as a symptom of 'bad' design for the same reason.

The proximity of elements in a graphic layout signal their relationship to the reader, e.g. pictures and their captions are usually adjacent. A design which ignores the need to unite similar kinds of information or separate others which are different may be difficult to read. Furthermore where the space between words in lines of text is wider than the distance between successive lines then reading becomes more difficult than if the reverse were true.

A layout that exhibits any specified 'bad' symptoms can be defined as being a bad design. The violation of any of the 'bad symptom' rules in a rule-base would indicate the presence of a design defect.

However compliance with a rule-base does not indicate the absence of any other condition, nor combination of conditions. In medical diagnosis the absence of prescribed symptoms indicates that the patient does not have the condition(s) addressed by the rules, but may have another complaint which is not addressed by the system. Similarly, symptoms of bad design are an indicator that a design is bad, especially in a system where rules are applied consecutively, but the absence of prescribed design faults would be indicative only that the design is not bad according to a prescribed set of rules; other conditions may be present which are not prescribed.
A 'bad' graphic layout will therefore be defined as one which exhibits any specified 'bad' symptoms. Conversely, a graphic layout will be described as 'not bad' if no 'bad' symptoms are witnessed, i.e. none of the contingent rules of bad design are violated.

3.7 We can define 'not bad' designs

Consequently it is argued that aesthetic merit may be determined by the presence or absence of prescribed graphic design defects, i.e. an aesthetic graphic layout may be one which exhibits no bad symptoms. Therefore in this study it is proposed that a layout which exhibits no bad symptoms can be defined to be 'not bad', but this does not imply that it is 'good'.

It is my contention that the facility to declare that a design is not bad is

a. an improvement on the present situation where no attempt is made to provide aesthetic evaluation of this kind.

b. It is better than nothing at all because it provides information useful to the user in improving the composition of a layout.

c. It will be bettered only when universal consensus is attainable in the graphic design domain.

3.8 Conclusion

Graphic designers like other designers employ grids to provide consistency where more than one page layout is being composed. Grids are used to provide structure in a graphic layout.
The availability of new technology like computer graphics and Desktop Publishing systems provides the opportunity for new approaches and different solutions with which to oil the wheel of originality. To quote from a television broadcast which took as its theme graphic design; ‘as the possibilities increase, the role of design in the clear presentation of information becomes ever more important’, Garrett (1989).

Highly organized configurations rely on the use of a grid. The structure of the graphic layout serves to regulate the sequence, pattern and order of the information. The choice of starting point may determine the overall visual pattern and regulate the variety of possible options. This can be likened to a game of ‘noughts & crosses’ in which the first move often determines the outcome.

There are any number of reasons which will determine the starting point; e.g. the content and the theme of a graphic layout for example. There are however certain similarities in the visual characteristics of a graphic layout.

An experienced designer will attend to the distribution, alignment and geometric configuration in a composition. The experienced designer is equally concerned with the significance of the part that each feature will play in the communication and with attempting to empathise with the intended audience attempting to match their perceptions.
The things we like are sometimes described as possessing an aesthetic quality or aesthetic appeal. We may describe visual appearance as being aesthetically pleasing yet its analysis is amorphous, its presence is sometimes ethereal, difficult to articulate. Our capacity to articulate the reasons for our preferences is limited by an imprecise and ambiguous vocabulary, e.g. a glass which is half full may be described as being half empty.

Judging from the literature, expert agreement on good design is often unattainable, usually controversial and therefore inconclusive. It is here suggested that the presence of bad symptom(s) in a layout is a reliable indication that the design of a layout is not good. Conversely, the absence of bad symptoms indicates that a layout is not bad. Therefore it is proposed that a layout which contains no bad symptoms is declared to be 'not bad'.
CHAPTER 4: RULE-BASED SYSTEMS

4.1 Introduction

Artificial Intelligence (AI) is the generic name for the branch of computer science which includes expert systems, knowledge-based, and rule-based systems: some authors use these terms interchangeably, Diaper (1989). As the focus of this chapter is on rule-based systems, the term rule-based system is used except where it is necessary to highlight aspects of a specific system.

According to Born (1987) the term artificial intelligence was coined in 1956 by John McCarthy and Marvin Minsky. AI is an important part of computer science concerned with designing intelligent computer systems. An intelligent computer system is one which exhibits characteristics which we associate with intelligent human behaviour, e.g. understanding language, learning, reasoning, and problem solving, see Born (1987). Diaper (1989) defines intelligence as applied to machines as 'any machine that is able to deal with inputs that are incompletely specified ... from the user or the programmer'.

A rule-based system contains knowledge about a particular domain. Hayes-Roth et al (1983) report that 'knowledge in any speciality is usually of two sorts: public and private. Public knowledge includes the published definitions, facts, and theories of which textbooks and references in the domain
of study are typically composed.' And private knowledge 'consists largely of rules of thumb that have come to be called heuristics. Heuristics enable the human expert to make educated guesses when necessary, to recognize promising approaches to problems, and to deal effectively with errorful or incomplete data' like that encountered in a layout. A central task of building a rule-based system is eliciting such knowledge. There are several reasons for focussing on 'knowledge rather than formal reasoning methods'. Two of these reasons are:

i. 'human experts achieve outstanding results because they are knowledgeable. If computer programs embody and use this knowledge, then, they too should attain high levels of performance.'

ii. 'focussing on knowledge recognizes its intrinsic value.'

Rule-based systems are used for a variety of applications, e.g. interpretation, prediction, design configuration, planning systems, monitoring, debugging, repair, instruction, control and diagnosis. A diagnostic system relates observed irregularities with possible causes, see Diaper (1989).

4.2 The architecture of a typical rule-based system

A rule-based system contains knowledge acquired from experts in a specific domain which can be made available to non-expert users. We all possess knowledge which enables us to carry out everyday procedures from tying a
shoe lace, baking a cake, to the highly sophisticated diagnosis of human diseases. A doctor, a mechanic, and a graphic designer will have specific knowledge which will enable them to do their job; specific knowledge about their domain or area of specialism, i.e. domain expertise.

The process of obtaining this domain knowledge for a rule-based system is known as knowledge acquisition. Some domain knowledge is concerned with the explicit procedures used by the expert, e.g. the process used by a doctor in performing a diagnosis. Other domain knowledge is concerned with the implicit unobservable strategies by which the expert arrives at his diagnosis and goes on to provide a prognosis. The process of eliciting the expert's intuitive knowledge is known as knowledge elicitation.

Knowledge acquisition and knowledge elicitation are highly specialised techniques performed by knowledge engineers. The knowledge engineer may employ a variety of powerful data gathering techniques depending on the domain, intended knowledge application and rule-base being developed.

Some rule-based systems contain a knowledge acquisition module, most contain an explanation module to provide the user with an explanation of the line of reasoning used to arrive at a conclusion; all will have a 'knowledge base', inference engine and a user 'interface', (see Figure 4.1: Hopgood's model of a rule-based system).

97
A knowledge base contains knowledge expressed in the form of facts and rules e.g.

**Fact 1.1**
Joe Bloggs works for ACME

**Rule 1.1**
IF ?x works for ACME Then ?x earns a large salary

The syntax used may vary from system to system. This type of IF ... THEN rule is called a *production rule*, and takes the form If condition Then result, i.e.

IF the condition is met, THEN the result is invoked.
The declaration that something is TRUE or FALSE is known as a *proposition* or a *predicate*: TRUE (or FALSE) is known as the truth value. A proposition could be stated as

IF pedestrian light is green is TRUE

THEN the pedestrian can go is TRUE

If something is TRUE this is known as a *fact*.

The *knowledge base* contains predicates and their truth values. The left hand side of a rule is known as the *antecedent* and the right hand side as the *consequent*, i.e.

<table>
<thead>
<tr>
<th>Antecedent</th>
<th>Consequent</th>
</tr>
</thead>
<tbody>
<tr>
<td>IF ?x works for ACME</td>
<td>THEN ?x earns a large salary</td>
</tr>
</tbody>
</table>

IF the antecedent matches database, THEN the consequent is added to database.

Facts may be expressed in the form

IF P is TRUE, THEN q is true.

According to Hopgood (1993) facts may be made available to the system at the outset: these are known as *static facts*. Facts revealed while the system is running are known as *transient facts*. Both static and transient facts are described as 'given' facts.

One or more given facts may satisfy the antecedent of a rule, resulting in the generation of a new fact, known as a *derived fact*. For example, by applying rule 1.1 to fact 1.1, we can derive:
Fact 1.2
Joe Bloggs earns a large salary

The derived fact may satisfy, or partially satisfy, another rule, such as

Rule 1.2
IF ?x earns a large salary OR has job satisfaction
THEN ?x is professionally content.

... from which the system would deduce that Joe Bloggs is professionally content.

Therefore RULES 1.1 and 1.2 are interdependent since the antecedent of the first rule can satisfy the antecedent of the second.

4.2.1 Triggering and firing
If the antecedent truth value of a rule matches the truth values of the facts in the Fact Database it is said to be triggered. If the rule is then selected it is said to have fired. The antecedent of a rule must be satisfied before it fires. When a rule fires its consequent facts are added to or updated in the Fact Database. It is possible that more than one rule could be triggered, and these form the conflict set. The inference engine has to apply one of a variety of strategies for selecting which rule to fire from the conflict set. The conflict set has to be resolved so that one rule will fire.

4.2.2 Conflict resolution
The following are some examples of the strategies which might be applied to select which rule in the conflict set should fire:
First-come First-served:
In this strategy the first rule that is triggered is fired. This avoids the need to create a conflict set and is also fast. However it has the disadvantage that the first rule that fires may be the least important.

Prioritising the rules (rule ordering):
The rules are ranked and the most important rule is triggered first and fired.

Generality ordering:
The rule that has the most conditions to be satisfied is fired. For any two rules in the conflict which are triggered the one with the most conditions will be selected. This approach is also called size ordering or specificity ordering.

Context limiting:
Group the rules so that in any one cycle one group will be active thus reducing the chance of a conflict arising.

The consequence of a particular rule is added to the database and the cycle starts again.

4.2.3 Control mechanisms
According to Hopgood (1993) inferencing programs are known as inference engines. 'Inference engines vary greatly according to the type and complexity of the knowledge with which they deal. Two important types of
inference mechanisms can be distinguished, namely forward chaining and backward chaining. These may also be known as data driven and goal driven respectively', see Hopgood (1993).

4.2.4 Forward chaining

Forward chaining is known as data driven inference because the reasoning process starts by addressing the facts and rules and reasons forward, testing each rule and firing every rule whose antecedent is shown to be TRUE.

Inferencing in the forward chaining mode follows a series of cycles where the rules are tested and truth values are ascertained. During the initial inferencing cycle new data are generated. In the next cycle the rules are again tested. This process is performed until no more rules fire, figure 4.2. The outcome is unpredictable resulting in either the generation of new data or trivial and
irrelevant information. The forward chaining (data-driven) approach is useful for addressing problems of interpretation where the user wishes to know everything the system can report without referring to the user for new data. So in this approach the system will continue until all possible values are obtained. As shown in the example control is not always consecutive.

4.2.5 Backward chaining

Backward chaining is also known as goal driven due to the solution focussed strategy which applies. In backward chaining the existence of a solution (or goal) is assumed; the system tries to find evidence for the corresponding truth value of a proposition., figure 4.3. Using this strategy it is conceivable that the system may conclude that the data provided is insufficient to draw a conclusion.

In backward chaining the goal prescribed by the user appears in the consequent of a rule. The procedure begins by searching the knowledge base for rules which might give the required conclusion. Next a search of the rules antecedents is made to see what is needed to fire the rule. If the rule fires then the goal is proved. If the rules do not fire the triggered rule is 'stacked' and the user may be requested to provide a further value (new data) for a sub-goal to direct the continuing line of reasoning.

The backward chaining process can be either breadth-first or depth first. In breadth-first mode all the rules from which a value might be derived for the
The depth-first strategy is to backchain to the antecedent of the rule under consideration before moving on to any of the other rules. When the system is called upon, the inference module examines all the rules and facts to a prescribed forward chaining or backward chaining strategy. The knowledge base is updated when new values are derived or instantiated in the forward chaining mode and or when the user provides new values in the backward chaining mode.

4.2.6 Dialogue module / human computer interface or user interface
The dialogue module is sometimes referred to as the human computer interface or the user interface. According to Card, Moran and Newell (1983), ‘the key notion, is that the user and the computer engage in a communicative dialogue whose purpose is the accomplishment of some task. It can be termed a dialogue because both the computer and the user have access to the stream of
symbols flowing back and forth to accomplish the communication; each can interrupt, query, and correct the communication at various points in the process. All the mechanisms used in this dialogue constitute the interface: the physical devices, such as keyboards and displays, as well as computer's programs for controlling interaction.’ So human-computer interaction can be likened to a conversational dialogue albeit one based on unusual or contrived semantics.

Hart (1986) summarises this as ‘the input/output interface, which enables the user to supply facts and data and enables the system to ask questions or supply advice and explanation’.

Newman and Sproul (1979) report that ‘No single component of an interactive computer program is more unpredictable in performance than the user interface, i.e. the part of the program that determines how the user and the computer communicate’. The user interface has four components: the user’s model, the command language or dialogue, feedback and the information display.

The user’s model
Newman and Sproul explain that the user’s model is a mental model formed of the information he or she manipulates and the processes which he applies to the information. This mental model can be likened to grammar of a natural language. Fluency is achieved when the user is no longer conscious of applying the grammar. ‘Without this model the user can do little more than blindly follow instructions, like an inexperienced cook following a recipe’. In
some applications the user's model simulates a real world system. Often it is necessary to devise systems which do not resemble real world events, where processes are abstracted and simplified to facilitate ease of manipulation. The use of processes and concepts familiar to the user makes the use of the system and the user's mental model easier to learn and gain acceptance. To use familiar processes with different outcomes will create confusion.

Newman and Sproul represent the user's model as a set of objects and a set of actions the user applies to the objects. The user can control the display of the object and its movement, ask questions about it, modify it or erase it or replace it. According to Newman and Sproul 'the simplest object is a single character, collection of characters form words, words form lines, lines form paragraphs, paragraphs form pages, and a set of pages forms a complete document'. Actions may be confined to a single object or may apply to a class of objects.

The user's model contains intrinsic objects and control objects. Newman and Sproul distinguish intrinsic objects as 'those which are intrinsic to the application' and control objects as 'those whose purpose is to assist in the control of the program'. Examples of control objects include cursors and command menus.

The command language

When the user has gained an understanding of the user model either by explanation or demonstration he needs to understand the command language of
the system, i.e. the commands which must be used to initiate a procedure or respond to questions prompted by the computer program.

A command language which is simple to use is beneficial to both experienced programmers and inexperienced users.

Feedback
Feedback assists the user to operate the computer program. Feedback comes in the form of acknowledgement, query, explanation and confirmation by the repetition of a user entered data.

Information display
Typically, human computer interaction depends on information being presented to the user as a form of visual display. The information displayed tells the user about the effect of the commands issued by the user. The information shows the user the ‘state of the information he is manipulating’. The information displayed may be graphical or alphanumeric or both.

While there is a need to design these components independently, Newman and Sproul emphasise their interdependence in the computer system.

Computer/user dialogue for an expert system
Hopgood (1993) explains the computer/user dialogue as follows ‘Typically the user of a ‘rule-based’ system will enter into a dialogue in which he or she
describes the problem (such as the symptoms of a fault) and the ‘rule-based’ system offers advice, suggestions or recommendations. The dialogue may be led by the ‘rule-based’ system, so that the user responds to a series of questions or enters information into a spread sheet.

Alternatively, the ‘rule-based’ system may allow the user to take the initiative in the consultation by allowing him or her to supply information without necessarily being asked for it.

4.3 Advantages and disadvantages of rule-based systems

According to Hart (1986), there are various reasons for using a rule-based system. It would always be available. Several copies could be made for use in different locations and simultaneously. Its performance would be consistent and mistakes would be rare.

The expertise of two or more experts could be combined to provide information at least as good as any one of the individual experts from whom the information was acquired. Additionally the individual experts options could be offered to the user. Another reason is that they are particularly suited to undertake tedious and routine tasks.

Like Hart, Moore and Miles (1991) highlight the benefits of using more than one source of expertise but acknowledge the difficulties of combining the expertise of two or more experts. Furthermore the involvement of individuals of different status and length of experience is more likely to reveal
exceptions. Although the involvement of two or more individuals increases the time spent in interviewing ‘the quality of knowledge acquired was much richer’. They suggest that the advantages ‘can be epitomized in the amount of project time saved and the elicitation of a more complete knowledge base’.

However, the performance of expert systems based on the expertise of two or more experts is sometimes worse than that of systems based on one source of expertise because the addition of more than one opinion may introduce inconsistency (see, for example, Hopgood, 1993)

4.4 Difficulties in using rule-based systems

Hart (1986) suggests the following difficulties in the use of rule-based systems:

*Choice of domain* - Rule-based systems suffer from a variety of performance limitations. They do not perform well in domains where commonsense concepts apply. Their performance is poor in domains where a lack of consensus exist.

*Acceptability* - Many domains have a resistance to the use of automated techniques. Some experts are sceptical about their use and may continue to be so even when the system’s performance is in agreement with the experts.

*Uncertainty* - Rule-based systems perform poorly where any of the information is missing or uncertain, unlike human experts who may cope intuitively with uncertainty.
Updating - Rule-based systems are not suited to situations where regular updating other than periodic inspection to insure relevance of the prevailing knowledge is required.

Limitations - ‘A human expert knows his limitations’. Rule-based systems ‘tend always to provide an answer and thus there is a tendency to over-diagnose. This can be problematic, and it should be stressed that they are most sensibly used as tools to assist rather than replace’ human expertise.

Testing - It is very difficult to fully test an expert system off-site. Systems often perform poorly on-site because the system developers are unaware of how the system should perform in operation.

Behaviour - ‘Although the aim of an expert system is that it should imitate a human expert, there are very few which actually do this. Dialogues are usually directed by the program, and explanations can often be difficult to understand. Consultations tend to be program driven and not user driven, and the user often has to suffer unnecessary explanations or output in order to obtain an answer.’

4.5 Use of rule-based systems for design evaluation

Expert systems of varying complexity are now available to perform a variety of design procedures including design evaluation. Oxman and Gero (1987) designed an expert system which can simulate the behaviour of a design expert
in the architectural domain. The knowledge contained in this system is based on the prevailing rules for the layout of a domestic environment. Here it is observed that according to Lauzzana and Pocock-Williams (1988), the architectural domain lends itself to formalised methods, because of the physical constraints involved in building and the discipline's highly developed symbolic methods of representation.

Similarly, the physical relationships in a graphic design layout are suitable for formal representation. Moreover procedures which have been devised could be adapted to the evaluation of a graphic representation. For example, Dengel (1989) describes 'an analysis procedure which is able to interpret the semantic meaning of document parts' suitable for the 'recognition of logical constituents of an office document'. Dengel's research informs the feasibility of the systematic analysis and evaluation of electronically produced publications such as word processed documents and desktop publications.

Therefore it is postulated that a graphic design layout produced by a designer using a computer system could be systematically examined using similar rule-based techniques. Details could be compared with specified design criteria and deviation from ideal combinatorial possibilities noted and reported for the designer's reaction. In this way a computer system could assist the designer with complex computations and in routinely ensuring that a design specification is adhered to.
It is in the nature of the design process that designing necessitates continuous (re)evaluation of the item being evolved, from conception to completion. Typically in graphic design, the process of design evaluation is manifest in the graphic designer's recursive inspection of the elements in a layout for known bad symptoms from preliminary abstractions to completion. The designer's reaction is driven by both emotion and rationale. This thesis shows that some of the rationale could be systematically addressed in a computer system. Therefore a computer system could offer an effective method of rational design evaluation.

4.6 Arguments for and against the validity of expert systems

We can only replicate things in an expert system which can be understood objectively. Mechanical steps and the systematic processes used in some mathematic computations and engineering procedures, for example, can be replicated in computer systems. Proponents of expert systems suggest that our reasoning processes can be replicated in this way. For instance Lansdown (1982) suggests that the behaviour of human experts and their cognitive processes can be compared with conventional computer programs.

Opponents argue that computers, unlike minds, do not and cannot possess intentionality. But intention is not necessary or desirable in all human activity. In aesthetic evaluation for example, intention is a barrier to the necessary passive contemplation of the beautiful.
Furthermore Dreyfus and Dreyfus (1986) contend that human intelligence is a dynamic process employing both analysis and intuition and that this combination of human expertise cannot be fully elicited by knowledge engineers or captured in the form of rules. This acknowledges that presently it is possible to elicit some rules. The psychology of choice informs us that we do not make decisions by rules but by comparing alternatives. But it is evident that the comparison of alternatives is better performed objectively and consistently; neither of which humans are renowned for as Lansdown (1982) points out.

Like Lansdown, other researchers speculate on the benefits of utilizing automated checking and evaluation procedures in the engineering and architectural domain. Oxman and Gero (1987), and others provide evidence to support the efficacy of producing an automated system for design synthesis, diagnosis and design analysis and evaluation in the design domain and in painting.

Kirsch and Kirsch (1988) have devised experimental procedures for the systematic analysis of the anatomy of painting styles based on shape analysis. In particular they report experiments using shape grammar procedures to describe a visual representation. They demonstrate the degree of geometric complexity permitted by serial shape transposition and that subtle irregular curvilinear shapes can also be addressed. The suitability of a shape grammar analysis procedure devised for the analysis and synthesis of fine art paintings
supports the use of similar techniques for the aesthetic evaluation of a graphic layout.

Furthermore, the availability of computer systems able to analyse, compose, decompose and generate new visual representations according to the rules of a predefined prescription underline the computer’s potential for graphic design evaluation.

Rule based systems are programs which allow some human knowledge to be made explicit as facts and rules.
CHAPTER 5: RULES DERIVED FROM A PRIOR SOURCE

5.1 Introduction

The construction of a rule-based system relies on the availability of human knowledge and tools suitable for its elicitation. Knowledge elicitation is a problem in building a rule base because of the relative inaccessibility of human expertise. In this chapter the utilisation of public knowledge (Rubinstein's recommendations) representative of the knowledge generally available in specialist textbooks, journals and academic references is examined. Empirical tests of the data obtained in the elicitation of knowledge from this source will be used in chapter 7 to inform the elicitation of implicit (private) knowledge from graphic design experts. Examples of relevant knowledge elicitation procedures are reported.

5.2 Knowledge acquisition and elicitation

The construction of an expert system relies on the availability of expert human knowledge and a reliable method of knowledge acquisition and knowledge elicitation; the process of getting knowledge from experts for that domain. Diaper (1989) distinguishes knowledge acquisition as knowledge needed by the system and knowledge elicitation as the process of getting knowledge from a source, e.g. human expert(s). He suggests that the process of knowledge acquisition is concerned with getting the knowledge elicited from the expert
into the computer, i.e. knowledge representation which involves writing the knowledge in the language required by the expert system shell. Knowledge elicitation involves initial analysis to decide what knowledge is needed. According to Hayse-Roth et al (1983) the stages in knowledge acquisition can be characterized as problem identification, conceptualization, formalization, implementation and testing.

Knowledge elicitation is the process of extracting knowledge from a source, e.g. from references (documents and artefacts) or from human experts or even non-experts directly. It is the process by which facts, rules, heuristics and procedures used by human experts to solve problems in the particular domain are elicited for use in the building of rule-based systems.

Diaper (1989) lists some relevant methods of knowledge elicitation as follows:

**Interviewing**

This involves the interviewer asking the expert questions for later analysis.

**Focussed talk**

This technique 'sets the knowledge provider a verbalisation task and the knowledge provider responds by producing a verbal report and perhaps artifacts such as diagrams, graphs, or sketches'.

116
Teachback

The process serves to remind the knowledge provider of the information previously provided. This provides the opportunity for clarification by both the knowledge elicitor and the knowledge provider.

Critiquing

This approach is particularly useful in domains where there are alternative ways of achieving a result or there is no particular right solution. The process of critiquing requires the knowledge provider to identify positive or negative aspects, e.g. to point out good or bad features which can be noted.

Protocol analysis

This technique provides the scope to elicit a variety of visual and verbal data about the process involved in completing a task and the environment in which it is being undertaken. In the context of knowledge elicitation, protocols are understood to be external records made on audio and or video tape when a knowledge provider is engaged in the performance of a predetermined task; written notes may also be made. During this process the knowledge provider's actions can be readily observed. Additionally the knowledge provider is encouraged to report his or her thoughts. In this process the subject(s) are normally required to think aloud verbalising the thought processes behind their procedures. Think aloud and talk aloud are intended to give the elicitor an insight into the knowledge provider's strategies.
5.3 Knowledge elicitation for graphic design

Although graphic designers, like some other experts, find it difficult to articulate their implicit strategies, they communicate their expertise verbally and by demonstration: for example the apprentice approach to design practice and some forms of design education confirm the success of these methods of knowledge transmission according to Tunnicliffe (1991) and Schenk (1991).

Tunnicliffe (1991) offers an elicitation procedure devised for the elicitation of knowledge in the graphic design domain. In his opinion traditional methods of knowledge elicitation are inappropriate to graphic design because they offer a 'reductionist' view more suited to scientific problems. He contends that to reduce the overall problem of designing to discrete parts for analysis distorts the knowledge elicited and may miss essential design knowledge. But some reduction is possible, e.g. "margins" in posters.

Furthermore, he suggests that design knowledge elicitation requires a holistic approach not provided by scientific knowledge elicitation techniques. But holism and science are not mutually exclusive. He emphasises that graphic designing relies on narrative, metaphors, analogies and gestalts whereas science employs formal approaches associated with problem solving. In design knowledge elicitation he suggests, it is necessary to arrive at information relating to the interdependence of elements in a design, e.g. components such as 'Headline text', 'Body text', and 'Pictures' and 'weight, balance, colour, impact and emphasis'.

118
Tunnicliffe (1991) describes a study which he suggests demonstrates that a combination of protocol analysis and the teachback techniques offer useful tools for knowledge elicitation in the graphic design domain. He reports that these techniques allow the graphic designers to perform in a way natural to them. These procedures provide a holistic view of the graphic designer’s processes and facilitate the subsequent analysis of the verbal and visual data generated and recorded. He reports that the gathering of non-verbal (gesticular) information as captured in a video recording provides evidence of the dynamic nature of graphic designing. Together with a sound track, this can provide the ability to accurately retrace the process of designing through the development stages from conception to solution. Additionally the drawings and rough sketches produced provide primary data of the design discussions and support to the later ‘teachback’ sessions.

Of the techniques of knowledge acquisition and elicitation reported, critiquing is considered to be the most appropriate for this study. It is reported by Tunnicliffe that graphic designers routinely criticise their own work using explicit language and therefore should be happy to criticise the work of others. The identification of bad symptoms in a graphic layout does not require a protocol analysis. Critiquing is a relatively simple and inexpensive procedure to arrange and provides a rich source of data for analysis.

Knowledge elicitation in graphic design is difficult particularly because there is little consensus on the subtle features which contribute to the aesthetics of a visual representation and ambiguity in the application of a design prescription.
5.4 Rules extracted from a published source

There are evidently some canons of graphic design, as reported in chapter 3. One source of this information is in the design solutions generated by graphic designers. Analysis of the choice and relationship of elements in the content of a layout can reveal the preferences of the graphic designer e.g. the number of typefaces, type sizes, type setting, vertical and horizontal alignment and compositional symmetry or asymmetry are indicators of personal preference or adherence to traditional norms which can readily be determined by inspection. For instance, Tschichold (1991) reports the achievements of Van de Graaf, Rosarivo and Villard in revealing the key to the positioning of type areas from classically printed works of Guttenberg by inspection and configurational analysis.

Another source of the graphic designer's preferences is in the specialist literature in the graphic design domain. This source gives conjectural accounts from which a graphic designer's preferences can be determined. However there is evident inconsistency in the rules applied by professional graphic designers. Ruder (1981) and Spencer (1968) report conflicting views on justified type setting for example, which indicates the disparate preferences of two acknowledged authorities in this domain.

The literature on good design practice reveals that graphic designers' recommendations are frequently based on anecdotal evidence, and usually
stated in a mixture of intuitive phrases and specialist expressions which are too complex to be applied unambiguously in the systematic evaluation of graphic design layouts.

5.5 Rubinstein’s recommendations

Rubinstein (1988), a typographic researcher and consulting engineer, gives recommendations for graphic layouts which are stated in less ambiguous terms than those offered by most graphic designers. His work addresses the issue ‘How can the knowledge, art, and skill required for attractive and readable displays or printing be incorporated into the design of computer systems’. His aim is to illustrate ‘how traditional typography can and should be incorporated into computer system design, introducing computer professionals and engineers to the technical, aesthetic, and psychological principles underlying the appearance of characters, words, and whole pages’. The following are some of Rubinstein’s recommendations. These are quoted from illustrated comments extracted from a selection of detailed reports based on research of typographic conventions.

a. ‘The shape of text on a page should relate to the shape of the page.’

b. ‘Underlining is used with typewriters to indicate emphasis. In typesetting, use *italics* for in line textual emphasis and **bold** for greater emphasis, such as headings.’

c. ‘Set narrow lines ragged right’

d. ‘... for most books the main body of the text occupies only 50% to 60% of the page.’
e. ‘Typographic wisdom suggests that relatively few sizes, faces and variations are required. The number used should be limited to three.’

f. ‘Text with large interword spaces needs more leading. Large interword spaces often result when text is justified in narrow columns.’

5.5.1 Formulating rules from Rubinstein’s recommendations

Rubinstein’s recommendations were expressed in a mixture of domain specific terminology and natural language. My aim was to determine whether Rubinstein’s recommendations were unambiguous and meaningful. This required decomposing the recommendations to form related sub-sets of definitions as a prerequisite to the formation of coherent statements. My objective was to ensure that the rule(s) eventually formed retained their integrity in the formal interpretation, translation and representational reduction. Cognisant of the difficulty that ‘translation/communication, ... is never complete or errorless: it involves abuses ... swerves, mistakes, variations, differences’ (Johnson, 1988), a priority was to ensure that a reliable methodology of rule formulation was used to maintain the rules’ integrity.

The process of rule formulation underwent several iterations. Initially, Rubinstein’s recommendations were examined by the researcher to determine whether they could be understood; the researcher is an experienced graphic designer and therefore knowledgable in the methodology appropriate to this domain. From this examination the researcher perceived that the rules were
meaningful but that some should be simplified while others required elaboration to ensure the reader's comprehension.

Rubinstein's recommendations were then formally interpreted by the researcher and arranged in a list of simple statements which it was intended could be consistently and unambiguously applied to the page layout of a book. The following were the researcher's initial interpretation of Rubinstein's recommendations.

a. The text shape should match the page shape.

b. Do not underline.

c. Set narrow blocks of text of lines containing four words or less and four lines or more ranged left.

d. Text area should be less than 60% of page area.

e. There should be not more than three type sizes.

f. Interword horizontal space should be less than line space.

5.5.2 Can Rubinstein's rules be applied to books as intended?

Rubinstein aimed his recommendations at the designers of books. Preliminary tests in the application of these rules to Rubinstein's own book showed that he had adhered to his recommendations. To determine the reliability of these recommendations the researcher tested them in the evaluation of ten other books. The books evaluated were produced by different publishers, on various subjects ranging from a computer science text
book to a child's science fiction book. These books were of different formats and sizes.

Ten pages from each book were evaluated starting at page ten and thereafter every tenth page. In a few instances the tenth page was blank, e.g. because it coincided with the close of a chapter; then the eleventh page was evaluated. The evaluation was undertaken by the researcher who applied the individual recommendations on the list to each designated page sequentially. The details on the pages to which the recommendations related were examined by inspection.

Each recommendation was applied individually. In this test it was found that none of Rubinstein's recommendations was violated in the sample pages examined.

This result suggests that Rubinstein's recommendations were based on conventional values shared by the graphic designers who had prepared the books in the evaluation. It also showed that there were certain norms of book design to which their designers consistently adhered.

Although perhaps inconclusive, this test supported the belief that graphic design recommendations might be formulated into expressions which could be applied in a simulated systematic procedure of evaluation.
5.6 Summary

Various methods of knowledge elicitation have been reported in this chapter. Those particularly suited for the graphic design domain have been described. The utility of extracting graphic design recommendations from the literature has been examined. The main benefits of this source of knowledge is its accessibility and critiquing costs far less than protocol analysis.

A possible flaw in this approach is that of being unable to verify the rules interpreted and formulated by the researcher (knowledge engineer) with the rules originator. However the interpretational difficulties were minimised by the researchers knowledge of the graphic design domain. The results support the belief that recommendations such as Rubinstein's could be formulating into representational rules from the anglicised expressions devised here.
CHAPTER 6: THE EXPERT SYSTEM TEST

6.1 Introduction

In chapter 3, a variety of aesthetic issues of concern in the graphic design domain were discussed and rules of graphic design as described in several published sources were presented. Chapter 5 presented rules derived from one of these sources, i.e. Rubinstein's recommendations. The central issue in this chapter is whether these rules could be formulated into representational rules usable in a computer system.

Analysis of graphic designers' rules informed by their aesthetic values reveals that there is confusion in the subtlety of expressions used and inconsistency in the way graphic design rules are applied. The difficulty is whether the confusion which exists can be removed. One method of making the rules unequivocal is to determine whether they can be formulated into representational statements which can be entered into a computer program. In this procedure the computer program acts in arbitration of the rules' validity. The method used is known as the 'Expert System Test' (EST).

6.2 The formulation of representational rules from natural language

The Expert System Test (EST) is a methodology used to determine whether the natural language expressions used by experts can be formulated into explicit rules. The proposition of the EST is that rules which can be entered into a
computer system are explicit. The successful outcome of the EST would be confirmation that the rules elicited are unambiguous, logical, objective and could be consistently applied (Johnson, J. 1990a).

Several iterations were necessary before Rubinstein's recommendations could be formulated into logical expressions which were both understandable by non-designers, unequivocal in their application and ultimately usable in a rule base.

6.2.1 First iteration

It was decided to determine whether non-designers could understand Rubinstein's recommendations. Two academics without expertise in graphic design acted as subjects to assist with this process.

The subjects examined the recommendations to determine whether they could be understood by others outside the specialist domain of graphic design. This revealed uncertainty in the subjects' interpretation of the expressions used and their perceived application of the recommendations. A cycle of revision and representation was undertaken to improve the subjects' understanding of the recommendations. By the conclusion of the fifth revision much of the subjects' misunderstanding of the recommendations had been removed. The restated rules were as follows:

a. The shape formed by the enclosure of the contents should match the shape of the page.

b. Do not underline individual words or a series of words in a line of text.
c. Set blocks of text of lines containing less than four words or four lines or more ranged left.

d. The area formed by the enclosure of the contents of a page should be of a size which is less than 60% of the page area.

e. No more than three type sizes should be used.

f. The horizontal space between words in a block of two or more lines should be less than the vertical space between consecutive vertically adjacent lines.

However on presentation to the subjects the need for clarification of the domain specific terminology was evident necessitating further revision.

6.2.2 Second iteration

I refined the recommendations and provided verbal definitions where necessary to enhance their understanding by the two subjects but it was found that inconsistencies remained. Subsequently the rules formulated from Rubinstein’s recommendations were reinforced by associated written explanation to inform their application and on this occasion the recommendations could be applied more consistently. The list of rules was now as follows:

a. The text shape should match the page shape.
   *For example, when a page is portrait orientation then the shape made by enclosing the contents of that page should also be portrait.*

b. Do not underline.
   *The text should not be underlined to avoid the coincidence with or intersection of descenders.*
c. Set blocks of text of lines containing four words or less and four lines or more ranged left.

d. Text area should be less than 60% of page area. 
   *Text area is the area within which the contents of a page i.e. text and imagery, are arranged.*

e. There should be not more than three type sizes. 
   *The number of type sizes on a single page of a book should not be more than three.*

f. Interword horizontal space should be less than interline space. 
   *The space between a pair of words on the same horizontal line in a sentence should be less than the space between the base line (bottom of 'x' height) and mean line (top of 'x' height) of a pair of words on vertically adjacent consecutive lines in the same block.*

The subjects found this formulation to be more understandable and the rules' application more reliable. It was decided to adopt this formulation of the rules for a series of subsequent tests.

Another method of testing the rules' formulation was to attempt to apply them to other populations of publications.

### 6.2.3 Can Rubinstein's recommendations be applied to other populations of publications?

To test the generalizability of Rubinstein's recommendations to other populations samples of ten magazines and ten commercially produced leaflets were evaluated following the procedure used for the previously reported evaluation of books, see section 5.5.2.
The ten magazines evaluated were of various kinds; from a specialist monthly American publication on typography to a Sunday supplement, and were of different proportions, formats and sizes.

The magazines were selected on the basis of their heterogeneity in the population of magazine layouts. The selection was confined to magazines which had a large number of pages to ensure that a relatively large number of pages at a ten page interval were available for the evaluation. A total of seventy-six pages was evaluated.

The ten leaflets were drawn from different sources and themes from ceramics to a mailshot promoting tea. Seven were A4 size and three were of various smaller sizes; all were larger than A5.

The percentage figures shown in Table 6.1 refer to the number of times that each of Rubinstein's recommendations was violated in a total of seventy-six pages from the ten magazines and ten commercially produced leaflet samples. The variation in number of bad symptoms in the populations of magazines and commercially produced leaflets examined suggests that Rubinstein's recommendations for book layouts were not generalizable to other populations.

It is not possible to say whether different rules were applied to the design of the magazine page layouts evaluated than to the layout of books. It is observed for example that the size of margins in these classes of publications is different.
However it is evident that some of the other layout characteristics of magazines are different to those observed to apply to the design of the commercially produced leaflets.

This is concluded from the disagreement between the results obtained for recommendations ‘e’ and ‘f’.

<table>
<thead>
<tr>
<th>Rubinstein’s recommendations</th>
<th>Ten Magazines</th>
<th>Ten Commercially produced leaflets</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. The text shape should match the page shape.</td>
<td>26%</td>
<td>30%</td>
</tr>
<tr>
<td>b. Do not underline.</td>
<td>11%</td>
<td>10%</td>
</tr>
<tr>
<td>c. Set blocks of text of lines containing four words or less and four lines or more ranged left.</td>
<td>18%</td>
<td>20%</td>
</tr>
<tr>
<td>d. Text area should be less than 60% of page area.</td>
<td>37%</td>
<td>50%</td>
</tr>
<tr>
<td>e. There should be not more than three type sizes.</td>
<td>38%</td>
<td>70%</td>
</tr>
<tr>
<td>f. Interword horizontal space space should be less than interline space.</td>
<td>20%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 6.1: Rubinstein’s recommendations violated in magazines and commercially produced leaflet samples
6.2.4 Can the Rubinstein rules be applied to desktop publications of A4 posters and notices?

From a survey of the notices and posters displayed on notice boards at the Open University it was apparent that their layouts exhibited a richer source of bad design symptoms than publications produced by graphic design professionals.

This is very important from the viewpoint of desktop publishing (DTP) which is very attractive to those with minimal design training wishing to produce professional looking publications but possessing little knowledge of the norms of aesthetics which should be applied.

A preliminary assessment of the available cross section of notice board posters within the Open University suggested that a high proportion was produced by non-designers using DTP or word processor systems.

Forty A4 poster specimens judged to be produced by novices were selected from the population of notice board posters on display within the Open University, figure 6.1.

The researcher's evaluation of the forty DTP publications following the previous procedure, gave the results, (shown in Table 6.2) for the violation of Rubinstein's recommendations.
Figure 6.1: Six of the forty A4 poster specimens
Rubinstein’s recommendations

DTP publications

a. The text shape should match the page shape. 25%
b. Do not underline. 13%
c. Set blocks of text of lines containing four words or less and four lines or more ranged left. 18%
d. Text area should be less than 60% of page area. 23%
e. There should only be three type sizes. 33%
f. Interword horizontal space should be less than interline space. 23%

Table 6.2: Rubinstein’s recommendations violated in the forty A4 specimens

6.2.5 Experimental shortcomings and revisions

It was found that some of Rubinstein’s recommendations were difficult to apply, particularly those which relied on accurate measurement and calculation in comparing the area of the page with the area of the contents. Another problem related to the recommendation that ‘Text area should be less than 60% of page area.’ It was evident that this percentage figure was too low for magazines, where in some instances text occupied more than 80% of the page area. An inspection of the commercially produced leaflets revealed that more than 70% of the page area was occupied by text and non-text elements in these also.
This focussed attention on the margins and the variability of their ratios. Here attempts were made to formulate unambiguous statements which could address the relative uniformity of margins.

Further analysis of the specimens highlighted other difficulties associated with the compound effects of type legibility and readability which it was decided to address. In particular difficulties relating to the space between words and the proximity and alignment of text to imagery were identified.

It was evident that although the Rubinstein recommendations could be applied unambiguously to books there were difficulties in the application of his recommendations to the other populations of publications tested.

6.2.6 Can difficulties and ambiguities be removed?
To address these issues the original list of six recommendations was augmented by the addition of new recommendations and the removal of those which remained difficult to apply. In the modification some of the previous single recommendations were reformulated as two or more codependent statements and others were expanded for clarification. These formed the basis of a set of rules which were used in a test to determine their suitability for a simulated rule-based application.

6.2.7 Third iteration
The following list of twelve rules prepared for the simulation was compiled. These have been subdivided into four discrete groups to aid the reader's understanding of the nature of graphic design rules.
**Margins:**

1. If there are margins then left and right margins should be equal.

2. Top margin should be $\leq$ bottom margin.

**Alignment:**

3. Where blocks of type are in the same vertical column then any vertical separation on the left side should be $\geq$ the largest type size used in these blocks.

4. Where two blocks are horizontally adjacent but not horizontally aligned and the largest block is 7 lines or more then their separation should be $\geq 3 \times$ type size of the smaller block.

5. Where two adjacent blocks are centred than the horizontal separation of their centre lines should be $\leq 1$ mm.

**Symmetry:**

6. Where all the blocks of type are centred then any non-centred image block should have a block on the other side of the centre line such that a reflection of either block about their centre line intersects to produce a block with area greater than 90% of both blocks.

7. When more than 70% of blocks are centred then all blocks should be centred.

**Readability:**

8. Blocks of text composed of capital letters should be four lines or less.

9. Blocks of text should not be in contact.

10. Space between underscoring or box rule and type should be $\geq$ the thickness of the line used.

11. Underscoring should not intersect type.

12. Underscoring should be $\geq$ length of line of type.
6.2.8 Can these rules be used by other people?

To determine whether these rules could be used by other people the following procedure for diagnosing bad symptoms in a sample of A4 desktop produced publications was devised:

- Select sample of forty publications
- Prepare specimens for evaluation
- Two subjects independently evaluate sample using rules
- Subjects provide written comments identifying difficulties in applying rules
- Analyse results
- Discuss findings with subjects

At this stage of development, reliable, intuitive feedback was essential. To facilitate this, human resources were again employed, this time, in an elementary procedure designed to simulate the partial application of a robust, intelligent, computer program and a sophisticated image recognition facility.

Preparation of the sample

The component objects in the layouts were abstracted to simplify the examination of their contents and configuration, see figure 6.2. This follows an approach used by Dengel (1989) which treats a layout as a visual composition which can be disassembled into abstract constituents e.g. text and picture components. This procedure in Dengel's words, 'does not require any character recognition; instead the layout structure of a document and the geometric profile of its parts serve to identify the various logical objects.'
Figure 6.2: Specimen layout showing abstract constituents of text and picture elements.

Key:

- Text block - ranged left  
- Text block - centred type  
- Image block

138
The properties within blocks such as text characteristics, factors relating to font identification and size were recorded for examination. Other attributes examined are the number of typefaces and type sizes in a text block. By this method the area occupied by text in a predominant type setting and its impact on other text blocks in the layout was readily observable. Using this approach a hierarchy of blocks and their constituents was determined.

Classes of objects:

The blocks and their contents were assigned the following semantic descriptors appropriate to classes of objects within a population of DTP publications described as A4 posters.

a. BLOCKS
   Content (boundary within which all elements can be circumscribed)
   Text
   Picture
   Bleed

b. MARGINS
   Left margin
   Right margin
   Top margin
   Bottom margin

c. ALIGNMENT
   Horizontal
   Vertical
   Horizontal centre separation
   Horizontal adjacent blocks
   Vertical adjacent blocks
   Horizontal touching blocks
   Vertical touching blocks
The available sample of forty DTP A4 publications previously evaluated using the six Rubinstein recommendations was reused for this experiment.

The researcher prepared a chart showing the list of rules. In the preliminary examination of the sample the rules had not applied to six of the specimens. These specimens were eliminated from the sample at this phase of the study to minimise the preparation and evaluation time. The remaining thirty-four specimens were evaluated by the researcher who recorded the violation of rules as well as their application and non-application. The result showed that the rules could be applied consistently by the researcher to specimens. This data was retained for eventual comparison with the responses of two unbiased subjects. The two academics who had previously assisted again acted as subjects for the application of these rules.

Subsequently, two charts were prepared and duplicates were issued to each of the subjects. Where the application of rules required computation or relied on accurate measurement, e.g. rules 1 and 2 relating to the comparative height...
and width of margins and rules 10, 11 and 12 relating to the space between lines and length of underscoring respectively, the researcher completed these entries.

The partially completed charts were included with the abstracted representations of the poster samples and delivered to each of the subjects. The evaluation was undertaken independently and at different times and locations. The subjects were required to visually examine the features in the layouts to which the rules applied in a partial simulation of the actions of a sophisticated image recognition system. The requirements and method of evaluation were discussed and verbally agreed prior to the formal evaluation of the samples by the subjects. This was to form the basis of the strategy used for a questionnaire and the eventual implementation later reported.

The subjects were also required to note where in their opinion the chart entry was noticeably at variance with their observations. Each subject was also requested to record the time taken to complete the task. This provided an estimate of the approximate time design experts from outside the university might need to complete similar experiments. The subjects’ observations were analysed and comments on the difficulties they had experienced in applying the rules were noted.

The following table number 6.3, shows the agreement between the two subjects in the identification of rule violations in the sample layouts.
Agreement on bad symptoms identified by the subjects is shown as a fraction of the number of times the rules were applicable:

Table 6.3: A comparison of the consistency of application of Rubinstein's rules applied by two academics.

<table>
<thead>
<tr>
<th>Rule no.</th>
<th>Number of posters to which rule could be applied</th>
<th>Number of bad symptoms agreed by the academics</th>
<th>Number of bad symptoms disagreed by the academics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>34</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>4.</td>
<td>9</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5.</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>7</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>7.</td>
<td>15</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>8.</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>9.</td>
<td>23</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10.</td>
<td>12</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>11.</td>
<td>7</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>12.</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Number of times the use of the rules were agreed by the two academics = 18
Number of times the academics disagreed on the application of the rules = 5
* Values computed by the researcher
Observations:

The results of these independent evaluations suggest that design rules could be applied in a rule based system. There was good agreement between the results of both subjects. Rule ten was evidently the most unreliable to apply.

The results and comments were discussed in a follow-up meeting between the two subjects and the researcher. Both subjects were critical of several of the rules for four reasons:

a. The rules could not be applied to the specimens unequivocally.

b. Several of the rule statements were ambiguous.

c. Certain rules required qualification.

d. Some vocabulary was not defined unambiguously.

The subjects' comments include 'what is separation' and 'what is horizontally aligned.' In a few instances the subjects' application of the rules was also at variance with the researcher.

Consequently it was agreed that a glossary of the terms and the compilation of suitable definitions was essential to ensure that the rules could not be misinterpreted. In addition the rules were reviewed and where necessary revised to enhance their application.

Problems of interpretation affected the certainty and speed at which the details
within the layout could be evaluated. Other comments like 'does adjacent apply despite column organization' highlighted the need for conditions and exceptions to be added to rule statements.

Only when the three participants (the two subjects and the researcher) were satisfied with the rules formulation, were the rules accepted.

The information received supported the hypothesis that rules could be elicited from a prior source. The procedure demonstrated that it was possible to devise logical expressions to describe semantic relationships in a layout design. However there remained the need and justification for the rules elicited to be formulated into representational expressions and validated using a rule-based system. The Leonardo expert system shell was utilised for this procedure.

6.2.9 Fourth iteration: Testing the formulated rules using the Leonardo expert system.

Following the methodology of the EST attempts were made to formulate a sample of the individual rules into representational expressions suitable for use in the Leonardo expert system. The individual rules were formulated, tested and when verified, added to the knowledge base. The expert system would act in arbitration on the rules validity.

The approach taken by the researcher was to study a prescription given for the
building of a sample knowledge base for Leonardo and in particular modify the illustration of a procedure for a strategy adaptable for diagnoses.

So parallel with the process of rule formulation attempts were being made to understand the procedure for writing representational rules in the Leonardo command language. It was also necessary to become conversant with the structure of an object frame to which default values had to be added to ensure the relevant truth value options were represented. Attempts were also made to devise suitable 'query preface' expressions using natural language dialogue to request an appropriate value from the user to ensure that goals or sub-goals could be instantiated.

During the early stages when the technique of rule formulation was being explored, the rules devised were based on Rubinstein's recommendations after the second iteration, see section 6.2. Leonardo incorporates a rules check facility which serves to check the rule syntax identifying errors with the display of an error message and the rule displayed and the problem highlighted for the users attention.

Indeed the rule based system cannot be executed until all problems concerned with rule syntax have been addressed. However this is no guarantee that the system will perform predictably. Initially there were many difficulties and the process proceeded by trial and error; but eventually the following prototype rule base was produced.
The "RUB1" rule base

/* RUB1 design critique
/* This is a simple demonstrator
/* to advise the novice of bad features
/* in the page design of a book
seek design

/* Rules for %grid_area

if page_height > 1
and page_width > 1
and content_height > 1
and content_width > 1
then page_area = page_height * page_width;
content_area = content_height * content_width;
%grid_area = 100*content_area/page_area

if %grid_area > 60
Then the design is "bad"

/* Rules for page_format

if page_height > page_width
then page_format is vertical
if page_height < page_width
then page_format is horizontal

if content_height > content_width
then content_format is vertical
if content_height < content_width
then content_format is horizontal

if %grid area < 60
and page_format is vertical
and content_format is horizontal
then the design is "bad"

Rule formulated from Rubinstein's recommendations for comparison with the rule base rules

d. Text area should be less than 0% of page area.

a. The text shape should match the page shape.
if \% grid_area < 60
  and page_format is horizontal
  and content_format is vertical
  then the design is "bad"

/* Rules for underline */
if text is underlined
  then the design is "bad"

/* Rules for justification */
if characters_in_line <= 12
  and typesetting is justified
  then the design is "bad"

/* Rules for word spacing */
if type_size < horizontal_word_space
  then the design is "bad"

/* Rules for typefaces */
if typefaces > 3
  then the design is "bad"

/* Rules for type_size */
if type_sizes > 4
  then the design is "bad"

/* Rules for line space */
if vertical_line_space < horizontal_word_space
  then the design is "bad"
if horizontal_word_space < vertical_line_space
  then the consultation_is_complete is yes;
  the design is "not bad"

b. Do not underline.

c. Set blocks of text of lines containing four words or less and four lines or more ranged left.

g. 1 Interword horizontal space should be less than interline space.
   (This rule was sub-divided into two interdependent elements for clarification),
   see Rule for line space, g.2.

f. There should be not more than three typefaces

e. There should be not more than three type sizes.

g. 2 Interword horizontal space should be less than interline space.
The consequent of the last rule was devised to derive the conclusion *the design is “not bad”* when all the rules have been applied and none found to be violated.

The above ordering of the rules shows that they were rearranged in the rule base in terms of their interdependence. The rule statements devised illustrate the transformation which was necessary to create representational rules suitable for a computer system from natural language. The knowledge base was executed in the default mode.

In Leonardo the default mode is backward chaining with opportunistic forward chaining. The rule base was tested and operated with some difficulties due the over complexity of the rule statements. It was anticipated that with further simplification the difficulties would be removed.

Informed by this experience the researcher was later able to prepare a rule base from the rules finally agreed for evaluating desktop publications.

The following rules were formulated for the development of a more robust prototype rule-based system for the evaluation of desktop publications. It will be evident that the natural language rules now more closely resemble the representational rules devised for entry into the rule base.
"RUB2" rule base

/* RUB2 design critique */
/* This is a simple demonstrator */
/* to advise the novice of bad features */
/* in the layout of an A4 publication */

/* seek design */

/* Rules for margins */
if left_margin > right_margin
  then margin_difference = left_margin - right_margin

if margin_difference > 1
  then the design is "bad"

if margin_difference < -1
  then the design is "bad"

/* Rule for head margin */
if top_margin >= bottom_margin
  then the design is "bad"

/* Rule for vertical separation */
if blocks_of_type_in_vertical_column > 1
  and largest_type_size > 1
  and vertical_separation_of_adjacent_blocks < largest_type_size
  then the design is "bad"

/* Rules for vertical_difference */
if blocks_of_type_in_row > 1
  and type_size_in_first_block = type_size_in_second_block
  and type_size_in_first_block lines or more then their separation should be >= 3*
and top_lines_are_aligned is yes
and lines_in_deepest_block >7
and lines_in_smallest_block >1
then vertical_difference =
   lines_in_deepest_block
   - lines_in_smallest_block

if vertical_difference <3
then design is "bad"

/* Rules for centring blocks */

if blocks_in_vertical_column >1
and centre_of_first_block_from_left_of_page >1
then horizontal_separation = centre_of_first_block_from_left_of_page - centre_of_second_block_from_left_of_page

if horizontal_separation >1
then the design is "bad"
if horizontal_separation <= -1
then the design is "bad"

/* Rule where all centred blocks */

if number_of_blocks_of_type >1
and blocks_of_type_are_centred is yes
and number_of_image_blocks > 1
and image_blocks_are_centred is no
and image_blocks_are_horizontally_adjacent is yes

then area_of_largest_block = height_of_largest_block * width_of_largest_block;
area_of_smallest_block = height_of_smallest_block * width_of_smallest_block;
%area_of_blocks = area_of_largest_block + area_of_smallest_block;
if area_of_smallest_block < 90% area_of_blocks

5. Where two adjacent blocks are centred then the horizontal separation of their centre lines should be <= 1 mm.

6. Where all the blocks of type are centred then any non-centre image block should have a block on the other side of the centre line such that a reflection of either block about their centre line intersect to produce a block with area greater than 90% of both blocks.
largest block
and distance from centre of page to centre of smallest block + width of smallest block
* 2 * height of largest block
< 90% area of blocks

then the design is “bad”

/* Rules for number of % centred blocks */

if number_of_blocks > 1
and number_of_centred_blocks < 70% number_of_blocks
then the design is “bad”

/* Rules for capital letters */

if block_of_text is composed of capital letters is yes
and number_of_lines > 4
then the design is “bad”

/* Rules for test blocks in contact */

if number_of_blocks > 1
and rectangles containing text blocks are in contact is true
then the design is “bad”

/* Rules for underlining */

if underlining_is_used is yes
and space_between_underlining < thickness_line_used
then the design is “bad”

/* Rule for intersecting type */

if underlining_is_used is yes
and underlining_intersects_type is true

7. When more than 70% of blocks are centred then all blocks should be centred.

8. Blocks of text composed of capital letters should be four lines or less.

9. Blocks of text should not be in contact.

10. Space between underlining or box rule and type should be >= the thickness of the line used.

11. Underscoring should not intersect type.
then the design is “bad”

/* Rule for length of underlining */  

if underlining_is_used is yes  
and length_of_underlining < length_of_  
text_underlined  
then the design is “bad”

6.3 Conclusions

This partial system served to confirm the utility of the procedure. Moreover,  
the expert system test was shown to be an effective methodology for validating  
the rules.

12. Underscoring should be  
>= length of line of type.
CHAPTER 7: EMPIRICAL STUDIES - RULES ELICITED FROM GRAPHIC DESIGN EXPERTS

7.1 Introduction

This chapter reports a set of experiments in which graphic design rules were elicited from two graphic design experts. In these experiments no attempt is made to judge a design to be "good", but the rules are sufficient to diagnose a design as 'bad'. By definition, a design that is not diagnosed to be bad is 'not bad', so that all 'good' designs will be 'not bad' but some 'not bad' designs may not be good.

The experiments show that graphic design rules can be elicited from expert graphic designers in a sufficiently explicit way for them to become operational in a computerised expert system.

The experiments show that the graphic design experts do not apply their rules consistently, but when the rules are made explicit they agree entirely with the diagnosis of the rule based system.

These experiments therefore support the thesis that there are rules governing the aesthetics of a graphic design, as perceived by practising graphic designers, which can be implemented on computers. These rules may not guarantee 'good' design but may assist the novice to produce designs which are 'not bad'.
A variety of recommended graphic design rules was presented in chapter 3, revealing aesthetic and configurational issues which are evidently of concern to graphic designers. The fact that some graphic design rules exist suggests that graphic designers perhaps instinctively (and un-knowingly) use a form of intuitive judgement which can be elicited and formalised into explicit expressions suitable for systematic serial application in a rule-based procedure.

This chapter deals with the issue of making explicit the implicit judgement of graphic design professionals. It addresses the issues of the expression of intuitive, emotive and configurational reporting and the difficulties associated with the interpretation of the professional's diverse and esoteric comments. It reports a procedure of rule elicitation through the diagnosis and reporting (in a critique) by two professionals of bad design symptoms in a sample of A4 desktop publications.

Also examined are the effectiveness and reliability of the rules formulated from the professional's intuitive design critique. This focuses on the (in)consistency of the professionals' diagnoses of bad symptoms and on factors which might inhibit accurate diagnosis and consistent reporting. The predictability of the professional's judgement is statistically analysed.

A new procedure for the diagnosis of bad symptoms in a population of A4 desktop produced publications is presented. The effectiveness of graphic design rules elicited from two professional graphic designers is assessed. The
method of interpretation and formulation of rule statements is also described. The procedure for testing the rules is presented. The professionals' intuitive diagnosis of bad symptoms in a sample of A4 desktop publications, giving written comments, reported in natural vocabulary and specialist terminology is used as a basis for formalizing design rules.

The formalized rules were used for the diagnosis of bad symptoms in a second sample of desktop publications. The professionals subsequently diagnose bad symptoms in this sample. The expressions and bad symptoms reported by the professionals' are interpreted into formalised rules. The results of the professionals diagnoses of bad symptoms in the sample are compared with the serial application of the formalised rules used in the systematic procedure. The performance of the two procedures is analysed to determine the correspondence of the rule-based diagnoses with those of the professionals.

In the reporting of this study the term 'expert' is used for the two graphic design professionals from whose critiques two sets of rules were formalized for use in the implementation. The term 'user' is used for those individuals who might employ the rule-based system for the evaluation of bad symptoms in a layout. 'Novice(s)' is a term used to describe a group of inexperienced individuals for whom the rule-based system would provide intelligent support to facilitate their understanding of the evaluation criteria and diagnostic procedure employed in this form of design evaluation. 'Designer' and 'non-designer' are terms used to describe the persons who prepared the specimen
layouts in the sample of publications tested in this study.

7.2 Rules elicited from graphic design experts

A suitable rationale for aesthetic evaluation through the diagnosis of bad symptoms in a layout was reported in chapter 3. A reliable procedure for knowledge elicitation and the methodology for the formulation of logical rules executable in a knowledge base were described in chapter 6.

The result of the previous experiments suggested that rules can be used to evaluate graphic design layouts. The experiments described in chapter 5 were based on a set of rules formulated from what had been shown to be reliable recommendations reported by Rubinstein (1988). It was decided to determine whether a graphic designer's first hand reports of bad symptoms in a sample of DTP publications could be formulated into an equally reliable set of rules. The success of this experiment would rely on the availability and co-operation of suitable graphic design experts.

Two graphic design experts agreed to take part in the proposed experiments. They will be referred to as (A) and (B) to preserve confidentiality. Both were previously known to the researcher. Their design experiences were complementary: (A) had undergone no formal tuition in graphic design, having acquired his experience through an apprenticeship, more than thirty years of professional experience working in the capacity of Art Editor for an international publishing company after taking degrees in English and Art
History and also a lecturer in art and graphic design for more than twenty years. (B) is a graphic designer and illustrator/print maker who received a conventional, formal art school education prior to professional practice in a graphic design studio. Notably, of the two, only (B) had a working experience of desktop publishing and was familiar with the constraints applied in producing some pieces of work in the sample.

7.2.1 Preparation of the sample of forty posters and notices

In this experiment the available sample of forty DTP A4 publications used in the previous experiment was again used. The posters were photocopied sequentially during the same print run. The specimens were duplicated singly to be the same size as the original on A4 sheets. Each copied poster was numbered consecutively from 1 to 40 in the top left corner, on the blank reverse side of the page.

Forty questionnaires were distributed to each of the two experts, together with the forty prepared posters and a covering letter explaining what the experts were required to do.

The experts were requested to inspect the sample of posters identifying any 'bad symptoms' in the specimens. Comments on each 'bad symptom' were to be reported on an accompanying questionnaire. Their comments were expected to be their personal view on the symptoms reported.
The experts were asked to assess the relative importance of individual symptoms by assigning a weighting on a scale 1, 2 or 3 corresponding to their considered importance; 1 represents not very important, 2 important and 3 very important. Written comments on each of their findings was also encouraged to expose the constraints on the experts diagnosis. The level of detail was to be to the expert's discretion but would clearly be a factor in the classification and discrimination of reported symptoms.

The experts were also requested to record the time taken to examine each specimen. The experiment was expected to take approximately one hour; the experts were not informed of this in advance to avoid this being a control issue. In the event (A) recorded a total time of 95 minutes with one blank time entry, (B) took a total of 107 minutes with three blank time entries. The experts' failure to record all the time entries was perhaps due in part to the duration and monotony of the procedure and a decline in attentiveness. Psychologists agree that fatigue has a detrimental effect on attentiveness. For example, Child (1981) reports that mental tasks of two hours should be interrupted at half hour intervals by short periods of rest.

Both experts informally requested the researcher to give examples of the bad symptoms to be reported. The researcher reiterated the instructions to avoid leading the subjects to the identification of a particular problem type. It was agreed that further verification might affect the scientific integrity of the study.
However not giving the experts a full explanation of the purpose of the task and their contribution is seen in retrospect to have contributed to inconsistency and perhaps poor motivation. The experts were unaware that their comments would be eventually formulated into rules. Indeed they were not informed how the information they provided would be used.

Both experts employed a similar approach to identifying design defects in the specimens. Bad symptoms were identified on the specimens by a freely drawn line encircling the feature which was also referred to in the questionnaire. Where there was more than one symptom, each was identified by an alphanumeric code. The comments on the associated questionnaire was similarly coded.

7.2.2 Description by the experts of the ‘bad symptoms’ detected in each specimen of the sample

In their descriptions and comments on the bad symptoms identified the experts used a mixture of ambiguous expressions, specialist terminology, and emotive statements as expected. These comments were used by the researcher in a knowledge elicitation exercise to establish the implicit or explicit ‘rules’ that were being applied by each of the experts. The following describes the phases of the knowledge elicitation.

Data analysis began with the transcription and interpretation of the experts’ responses. Where necessary a response was recomposed into logical English statements embracing the essence of the expert’s individual comment.
Subsequently each statement was formalised to simulate (for each expert) a robust set of rules appropriate for a simple rule-based application. Reported symptoms, e.g. physical details, were examined for spatial conformity and deviation.

7.2.3 Interpretation of free format speech to form technically defined and rigorous rules

The researcher’s experience and familiarity with the conventions, terminology and phraseology of the graphic design domain enhanced the interpretation of the experts’ comments. The rules were initially formulated by the researcher who attempted to interpret the expert comments. The formulation of representational rules followed the methodology of the Expert System Test.

However it was recognised that the interpretation of idiosyncratic words and phrases is often variable between individuals. Specialist idiom can also be inconsistent at different times and in different situations. Where a mixture of colloquial and intuitive terminology is applied the dichotomy of this mixture is compounded. Accurate interpretation relies on the precise and logical definitions of specialist expressions.

Consequently the interpretation was verified by asking the experts who originated the comments received by the researcher to confirm the meaning of many of the terms received in their responses. The information received was used in formulating an experts’ glossary.
To assist the elicitation process a list of one hundred expressions used by the experts, those used in the elicitation and others in the literature giving recommendations by graphic designers, was prepared. The experts were requested to study this compilation and note their understanding of the terms used; incorporating additional terms to the list where necessary.

This provided a glossary of some of the expressions used in their descriptions together with those used by other experts. There was a high level of consistency in the expressions used by both experts.

The experts' glossary ensured the accurate and unambiguous interpretation of the expressions, specialist terms and unique and conventional phrases used in the various comments provided by the experts and their formulation as accurate rule statements. Moreover, precision in the definitions ensured that the correct symptom was addressed.

Another, more general glossary was also produced, the original phraseology of the rules did not result in their unique, logical and consistent application. It was found through tests involving a mathematician acting as the rule inferencing agent that some rules could be varied indiscriminately to apply to two or more related but subtly different conditions. Subsequent didactic procedures in rule formalisation refined the precision of their application. The resulting compilation of a set of unique descriptions ensured that the right symptom was successfully targeted.
The following is the general glossary of definitions devised to assist in the interpretation of the formalised rules.

*Glossary of terms:*

The following definitions apply to various terms employed in the rule statements. These examples relate to terms devised to describe semantic properties abstracted in the layouts. The glossary of the terms used by the experts was used to inform rule formulation. The following definitions are the result of repeated refinement. Even so it was found that many definitions remained conditional.

**BLOCKS**

*Content block* - The smallest rectangular block within which all the elements on the page are contained; excepting a page on which an element touches one or more of its edges (see bleed).

*Text block* - The smallest rectangular block containing a word or a string of words.

*Picture* - Any mark that is not text. Also any imported text e.g. text which is positioned diagonally. Additionally any artwork, e.g. a symbol or company name.

*Picture block* - A block which contains a picture.
Bleed page - A page on which a picture touches its edge.

MARGINS

Left margin - The space between the left edge of the page and the left edge of the content block.

Right margin - The space between the right edge of the page and the right edge of the content block.

Top margin - The space between the top edge of the page and the top of the content block.

Bottom margin - The space between the bottom edge of the page and the bottom of the content block.

Formulation of rules:

Rules were formulated by the researcher from the symptoms of bad design identified by the experts. The process of formulating a rule was iterative, involving an initial formulation, testing and perhaps one or more attempts at reformulation before eventual acceptance, see Figure 7.1.

Follow-up interviews with the experts were also used to ensure that an identified bad symptom was being correctly interpreted by the researcher.
Figure 7.1: The reiterating process of rule formulation.

A selection of bad symptoms identified comments, and with the importance rating assigned by the expert and the eventual rule statement formulated are shown below.

Importance Rating:
[1] - not very important
[2] - important
[3] - very important

Information relating to importance was not used explicitly in the experiments. However a comparison of this information shows that the importance of a bad design symptom can vary from one specimen to another. One may speculate that this variation is due to the presence of two or more bad symptoms, or physical dichotomies, e.g. where there are seven type sizes and the composition is balanced or unbalanced.
Another speculation is that emotive factors may contribute to inconsistency, e.g. the subject may not be sympathetic to the style of type setting.

The weighting of bad symptoms may be of value in a rule-based system. A quite sophisticated system is envisaged in which knowing whether a rule is very important or not very important can be useful in deciding whether a bad symptom can be ignored or at which level of frequency it becomes critical. At a rudimentary level the designer of such a system may decide that ‘very important’ scores three points whereas ‘not very important’ is worth only one point. In this case a score of at least three points may be necessary to declare the design of a layout as bad.

In another system this information might be imparted to the user for arbitration. Issues like this remain for future research.

The following are examples of the bad symptoms reported by the experts and the rules elicited from their comments.

**Examples of the (A) symptoms and rules:**

Symptom(s):
*Bad gap (word spacing).* \[1\] [weighting (low)]

Reason(s):
*Such unfortunate alignment, causing rivers, in which the words on the line below are closer than the words on the same line distract from their intended sense.*

Rule(s) Formulated:
LINE SPACE SHOULD BE GREATER THAN WORD SPACE

Symptom(s):
The easy use of space at the top of the page is not echoed in the graceless foot of the page.  [1/2]  [weighting (low) 1 - 2]

Not visually centred.  [1/2]

Reason(s):
One is left with a feeling that the elements are out of phase - this must be counterproductive.

Rule(s) Formulated:
i. IF THERE ARE TOP AND BOTTOM MARGINS THEN THE TOP MARGIN SHOULD BE LESS THAN OR EQUAL TO THE BOTTOM MARGIN

ii. WHERE TWO OR MORE VERTICALLY ADJACENT BLOCKS OF TYPE ARE CENTRED THEN THE HORIZONTAL SEPARATION OF THEIR CENTRES SHOULD BE LESS THAN 1 MM

Examples of the (B) symptoms and rules:

Symptom(s):
Far too close to the top of the page.  [3]

Reason(s):
'top' heavy / otherwise OK

Rule(s) Formulated:
i. NOT ALL BLOCKS SHOULD BE ABOVE THE CENTRE OF THE PAGE
Symptom(s): Underlining - not even complete in some cases.  [3]

Bad spelling

Reason(s):
It's so awful that I wouldn't read it.

Underlining gets caught up with descenders and makes it hard to read.

166
Rule(s) Formulated:

i. UNDERLINING SHOULD NOT INTERSECT DESCENDERS

ii. UNDERLINING SHOULD BE GREATER THAN OR EQUAL TO THE LENGTH OF THE LINE OF TYPE WHICH IS UNDERLINED

Symptom(s):
More space at top than bottom. [3]

Thrown together, some centred type, some ranged left, some ranged left and then set right. [3]

Crooked pasting up. [3]

Reason(s):
just a mess

Rule(s) Formulated:

i. IF THERE ARE TOP AND BOTTOM MARGINS THEN THE TOP MARGIN SHOULD BE LESS THAN OR EQUAL TO THE BOTTOM MARGIN

ii. WHEN MORE THAN 70% OF BLOCKS ARE CENTRED THEN ALL BLOCKS SHOULD BE CENTRED

iii. WHERE MORE THAN 70% OF BLOCKS ARE CENTRED THEN FOR ANY NON-CENTRED BLOCK ON ONE SIDE OF THE CENTRE LINE THERE SHOULD BE A BLOCK ON THE OPPOSITE SIDE SUCH THAT A REFLECTION OF EITHER BLOCK ABOUT THE CENTRE LINE INTERSECTS THE OTHER TO PRODUCE A BLOCK WITH AREA GREATER THAN 90% OF BOTH BLOCKS.

iv. BLOCKS SHOULD BE PARALLEL TO THE HORIZONTAL EDGES OF THE PAGE OR AT AN ANGLE GREATER THAN 10 DEGREES

Symptom(s):
DISCO - because it is in large italic it doesn't look as if it is centred. (maybe it isn't). [2]
Reason(s):
None

Rule(s) Formulated:

i. ITALIC TYPE SHOULD NOT BE CENTRED

ii. WHERE TWO OR MORE VERTICALLY ADJACENT BLOCKS OF TYPE ARE CENTRED THEN THE HORIZONTAL SEPARATION OF THEIR CENTRES SHOULD BE LESS THAN 1 MM.

From the expert’s comments on the symptom identified it was difficult to determine by observation whether two lines of italic text were centred. By measurement the difference between the centres was found to be greater than the permitted maximum of 1mm.

Comparison of the rules elicited in the test:

The (A) rules: elicited from Expert A

a1. Blocks should be vertically aligned. Similar to b24

a2. Text should not be set in caps.

a3. Text type should be equal to or less than 14 points.

a4. A block of capital letters should be less than four lines.

a5. Characters should not be bitmapped. Similar to b1

a6. Words in a line should be separated by the maximum of a lowercase ‘m’ in the fount used. Similar to b2

a7. The indentation of type in vertically adjacent blocks should be greater than or equal to the largest type size within the block.

a8. Where two or more vertically adjacent blocks of type are centred then the horizontal separation of their centres should be less than 1mm.
a9. If more than 40% of blocks are centred then all blocks should be centred.

a10. Sub-headings should not be extended from the edge of the descending block.

a11. Headings should be at the top of the page.

a12. A line of type at the top of the page should not be in the smallest type size.

a13. If there are top and bottom margins then the top margin should be less than or equal to the bottom margin. *Similar to b5*

a14. Typeface used for title and main heading should be from the same family.

a15. Text of more than four lines should not be set solid.

a16. There should be a minimum of 12points (vertical) line space between blocks of type in different fonts.

a17. The space between words on the same line in a block should be less than the space between lines.

a18. There should be less than five type faces. *Similar to b6*

a19. The space between vertically adjacent blocks of centred type should be less than 2.1/2 x the smallest type size in these blocks. *Similar to b7*

a20. The space separating a line of type and a line-rule underscoring it should be more than 2 x the thickness of the line rule.

a21. Where two or more vertically adjacent blocks of text are centred then the horizontal separation of there centres should be less than 1 mm.

a35. Bullet points should not appear to the right of a line of type.

a36. The space between blocks of text of the same fount should be less than the space between different founts.
a37. Underscoring should not intersect descenders.  
    Similar to b20

a38. An underscoring line should be greater than or equal to the length of the 
    line of type underscored. Similar to b21

a39. The space between sub-headings and block of text of the same fount 
    should be ranged left

a40. If text is ranged left then the other type should be ranged left.

a41. Within a box-rule the space at the top should be less than or equal to the 
    space at the bottom.

a42. The text type-size should be less than 14 points.

a43. Blocks of text should be set in a light or italic typeface.

a44. There should be a minimum of 12 points (vertical) line space between 
    blocks of type in different founts.

a45. If there are left and right margins then they should be equal.

a46. The space between blocks of text should be less than 4 x type size.

a47. Italic should be used to emphasize text within a block.

The (B) rules elicited from Expert B:

b1. Letters should not be bitmapped. Similar to a5

b2. The space between words should be less than the width of a lowercase 
    ‘n’ in the same fount. Similar to a6

b3. Headings should be heavier weight than text. Similar to a23

b4. Characters should be readable.

b5. If there are top and bottom margins then the top margin should be less 
    than or equal to the bottom margin. Similar to a13
b6. There should be less than 4 typefaces. Similar to a18

b7. The space between blocks of text of the same font should be less than 2. 1/2 x type size. Similar to a19

b8. Vertically adjacent characters should not form a word.

b9. The width of a line of type at the bottom of the page should be less than or equal to the width of lines of type at the top of the page.

b10. The type size of letters below the centre of the page should be less than type size of letters above the centre of the page.

b11. Drop initial should not touch letters in the line of adjacent text.

b12. The distance between the base line of type in consecutive lines of text should be greater than the type size of the text.

b13. Text should be in a lighter weight than the heading.

b14. The space between vertically adjacent blocks of text of different fonts should be less than 2. 1/2 x the smallest type size.

b15. The space between a block of text and the heading should be less than 2. 1/2 times the type size of the text.

b16. The weight of type below the centre of the page should be less than above the centre of the page.

b17. The length of lines of text below the centre of the page should be less than lines above the centre of the page.

b18. If there are blocks of text at the top and bottom of the page in the same font and setting, then they should be aligned.

b19. If all blocks are above the centre of the page then the design is bad.

b20. Underlining should not intersect descenders. Similar to a37

b21. Underlining should be greater than or equal to the line of type which is underlined. Similar to a38

171
b22. When more than 70% of blocks are centred then all blocks should be centred. Similar to a22

b23. Where more than 70% of blocks are centred then for any non-centred block on one side of the centre line there should be a block on the opposite side such that a reflection of either block about the centre line intersects the other to produce a block with area greater than 90% of both blocks.

b24. Blocks should be parallel to the horizontal edge of the page or at an angle greater than 10 degrees. Similar to a1

b25. If logo type is ranged left then it should be positioned left of the centre of the page. If logo type is ranged right it should be positioned right of centre of the page.

b26. The space occupied by logo and heading should be less than 1/4 of page area.

b27. If there are three or more vertically adjacent blocks of the same size font then the distance between the adjacent edges of any two blocks should be less than 2.1/2 x the type size in these blocks.

b28. The space between blocks of type should be greater than 1.1/2 but less than 3 x type size of the largest block.

b29. Italic type in a block of two or more lines of type should not be centred.

Although there was overlap of similar, and sometimes identical, rules used by the two experts, there were also substantial differences between the two rule sets.

Number of rules used by the experts:

The (A) rules = 47

The (B) rules = 29

172
Differences in the rules used by the experts:
Number of rules used by (A) alone = 37
Number of rules used by (B) alone = 19

Coincidence in the rules used by the experts:
Number of rules shared by (A) and (B) = 10

A variation of four of the augmented set of rules formulated previously by Glaze (re page 137) from Rubinstein’s recommendations were also found to be used by both of the graphic design experts; these rules are shown below.

a13 & b5  Top margin should be <= bottom margin.

a22 & b23  When more than 70% of blocks are centred then all blocks should be centred.

a37 & b21  Underscoring should not intersect type.

a38 & b22  Underscoring should be >= length of line of type.

This confirms that some of the rules formulated from Rubinstein’s recommendations are obeyed by some graphic designers consciously or subconsciously. It also suggests that some graphic design rules are rooted in a universally agreed, but perhaps unarticulated standard.

7.3 Testing the rules elicited from the designers
The previous experiment involved the assistance of two design experts in the diagnoses of bad symptoms in the sample of forty available DTP publications.
Subsequently, a set of rules was elicited from each expert's diagnosis. Following the methodology of the Expert System Test an experiment was devised to test the reliability of the rules.

The objective now was to compare the performance of a simulated rule-based system with that of the expert on a new set of publications from the same general population. The purpose was to determine the effectiveness of the rules and whether the expert's application of the rules could be predicted on the new sample of thirty posters.

7.3.1 Experimental design and method of analysis

Procedure:
a. Apply the rule-based diagnostic system to the new sample of posters unseen by the design experts.

A list of rules previously elicited from expert (A) was compiled, sifted and applied by the researcher to the new sample of thirty specimen poster layouts. The results were recorded (Table 7.1) as a prediction of the expert's evaluation. The unmarked samples were then sent to expert (A) for his diagnosis and return.

b. Obtain the expert's diagnoses of bad design symptoms on the new sample of poster layouts. The results are given in Table 7.1.

The same procedure was repeated in a test of the predictability of expert (B)'s rules, and the results are given in Table 7.2.

Rule sets:

Crucial to the experiment was the need to determine beforehand the most
appropriate statistical design for the data characteristics. Moreover it was essential that objectivity and repeatability were emphasised and that a rigorous experimental method was devised.

A study of the rule sets revealed that several applied insufficiently to be useful for the purpose of statistical analysis. Pilot studies of the statistical method had shown that rules which applied to fewer than 15 posters would generate unreliable results. Moreover, nothing could be concluded where a rule applied to fewer than 5 posters. Consequently several rules were eliminated because they failed to satisfy these pre-conditions; leaving 17 A and 13 B rules.

**Sample size:**

From a study of the statistical requirements (in section 7.4) it had emerged that thirty was the minimum number of specimens required to ensure statistically reliable results from the implementation. It was also essential to determine the minimum sample size to ensure that the experts' time was not used unnecessarily. Consequently a new set of thirty posters was collected following the procedure described in the previous experiment. These posters were presented to the experts for their evaluation. Meanwhile the rule base evaluation was done by me.

**7.3.2 Application of the rule sets to the 30 new posters**

The specimens were evaluated by me using the above rules. The predictions recorded are shown below. The figures show the number of times 'bad'
Table 7.1: Comparison of a rule based diagnosis using (A)'s rules, with his repeated diagnosis of bad symptoms in the same sample.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Number of posters to which the rule could be applied</th>
<th>Bad symptoms diagnosed by rule based system</th>
<th>Bad symptoms diagnosed by expert designer A, visit 1</th>
<th>Bad symptoms diagnosed by expert designer A with experiment</th>
<th>Bad symptoms diagnosed by expert designer A, visit 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1.</td>
<td>If there are left and right margins, then they should be equal; + or - 2 mm., excepting when a picture bleeds off the page.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>17</td>
<td>0</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>A2.</td>
<td>If there are top and bottom margins, then the top margin should be less than or equal to bottom margin.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>8</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>A3.</td>
<td>Blocks should not be in contact - vertically</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>A4.</td>
<td>Blocks should not be in contact - horizontally.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>A5.</td>
<td>Word blocks should not touch - horizontally.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A6.</td>
<td>Vertically adjacent blocks should be vertically left and right aligned.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>A7.</td>
<td>When more than 70% of blocks are centred then all blocks should be centred.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>9</td>
<td>4</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>A8.</td>
<td>Blocks of capital letters should be equal to or less than 4 lines.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>A9.</td>
<td>For three or more vertically adjacent text blocks of the same fount, the space between the tops and bottoms should be equal.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>A10.</td>
<td>The longest line in a non-list block of text greater than four lines should be a minimum of twenty characters.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A11.</td>
<td>There should be less than four typefaces.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>A12.</td>
<td>There should be less than five type sizes; excepting lists, tables and diagrams.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>A13.</td>
<td>The words in a line should be separated by less than or equal width to a lowercase 'n' in the fount used; excepting lists and tables.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>12</td>
<td>9</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>A14.</td>
<td>Characters in a word should not touch; excepting ligatures, script or italic text and imported artwork.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>A15.</td>
<td>Ascenders and descenders in consecutive lines of text should not touch.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A16.</td>
<td>The largest block must be above the horizontal centre of the page. If more than one large blocks of equal size, then at least one must obey this rule.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>11</td>
<td>2</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>A17.</td>
<td>The distance between centre lines of vertically adjacent blocks of centred strings of words must be less than 1 mm.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>23</td>
<td>1</td>
<td>23</td>
<td>9</td>
</tr>
</tbody>
</table>
symptoms were observed as a fraction of the number of times the rules were applicable in the sample of thirty posters. For example, for A1 there were 17 bad symptoms observed in 28 applications of expert A's first rule.

Table 7.2: Comparison of a rule based diagnosis using (B)'s rules, with her diagnosis of bad symptoms in the same sample.

<table>
<thead>
<tr>
<th></th>
<th>Number of posters to which the rule could be applied</th>
<th>Bad symptoms diagnosed by rule based system</th>
<th>Bad symptoms diagnosed by expert B</th>
<th>Bad symptoms diagnosed by expert B with experimenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>If there are left and right margins, then they should be equal; + or - 2 mm.</td>
<td>28</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>B2</td>
<td>If there are top and bottom margins, then the top margin should be less than or equal to bottom margin.</td>
<td>28</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>B3</td>
<td>Characters should be recognisable.</td>
<td>30</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>B4</td>
<td>Characters should not touch.</td>
<td>30</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>B5</td>
<td>The largest block must be above the horizontal centre of the page. If more than one large block of equal size then at least one must obey this rule.</td>
<td>30</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>B6</td>
<td>There should be less than four typefaces; excepting any imported artworks.</td>
<td>30</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>B7</td>
<td>If there are three or more vertically adjacent blocks of the same font then the distance between the adjacent edges of any two of these blocks should be less than (2 \times \frac{1}{2} \times \text{type size in these blocks.} )</td>
<td>15</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>B8</td>
<td>Vertically adjacent characters should not form a word.</td>
<td>30</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B9</td>
<td>The width of blocks below the centre of the page should be less than at least one block above the centre of the page.</td>
<td>30</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>B10</td>
<td>The type size of text below the centre of the page should be less than the largest type size above the centre of the page.</td>
<td>30</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>B11</td>
<td>The bottom margin should be less than (\frac{1}{3}) of the page height.</td>
<td>30</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B12</td>
<td>Within a box rule the space at the top the top should be less than or equal to the space at the bottom.</td>
<td>15</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>B13</td>
<td>Blocks should be parallel to the horizontal edges of the page or at an angle greater than 10 degrees from the horizontal.</td>
<td>30</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
7.3.3 The experts' evaluation of the posters

Following the same procedure as described in section 7.3.1, the samples discussed in the previous section were distributed to (A) and (B) for their independent observations and comments on the bad symptoms identified in the specimen layouts in the sample.

Comments recorded on the experts' questionnaires:
Separate lists of the experts' comments were made and later compared. It was noted that the degree of detail varied between the experts and from one specimen to the other as expected from the previous responses.

Often different symptoms in the same specimens were identified by the experts. Moreover designers describe the same symptom using personal expressions in a personal blend of specialist and emotive terms. They also attach different levels of importance to the same symptom. This suggests that the criteria experts employ are different and their values are variable.

I examined the returned samples and each of the experts' comments to determine which of the formalised rules applied. This procedure relied on the accurate translation and matching of rules formulated from the comments reported in the old with those reported in the diagnosis of the new samples.

A list of the bad features abstracted from the comments reported by each expert and the number of times the comments applied was recorded, see tables 7.1 and 7.2.
While there were forty specimens in the old sample there were just thirty specimens in the new sample. The general standard of design of the new sample of poster layouts was reported by the experts to be ‘much better’ than the design of layouts in the old sample. This was perhaps due to the growing awareness of the producers of the posters. As a result fewer bad symptoms were reported by the experts in the new sample.

In the experts’ responses to the new sample, subtle variations to previous comments were apparent and some new comments were reported.

As explained previously in section 7.3.1, the number of rules used by the researcher in the evaluation of the new sample of posters was:

The (A) rules = 17
The (B) rules = 13

Comparison of the rules used by the experts in this evaluation:

Number of rules used by (A) alone = 13
Number of rules used by (B) alone = 9

Coincidence in the rules used by the experts:

Number of rules shared by (A) and (B) = 4

7.4 Analysis of the data - Visit 1

Essential to the research was the statistical analysis of the data. A statistician had previously been consulted to determine the appropriate statistical approach for this experiment. He advised on the sample size and the statistical test after
determination of the experimental aim. Specifically he advised

- The Fisher exact probability test
- A sample size of thirty specimens
- The null hypothesis:

The null hypothesis states that the experimental results are due to chance, i.e. the experimental results are unpredictable. So we are looking for data to reject the null hypothesis and to show that our results are predictable.

The probability that our result is obtained by chance is 5% or 0.05. on the recommendation of the statistician. So where a figure less than 0.05 is computed from the distribution of frequencies in the statistical data then the null hypothesis can be rejected.

I collected the data required for the statistical analysis and examined it with the statistician. Analysis of the collated data gave the results shown at the end of this section.

Aim:
To determine if the experts tend to agree with the application of their rules.

Procedure:
Fisher's exact test (below) was used in the analysis of the data. This test is based on the association between the researcher's identification of bad symptoms by the systematic application of the expert's rules and the expert's
reported use of the same rules. The Fisher exact test is appropriate where figures
smaller than 5 occur in the data, see Siegel (1956). According to Siegel

'The Fisher exact probability test is an extremely useful nonparametric
technique for analyzing discrete data ... when the two independent samples are
small in size. It is used when the scores from two independent random
samples all fall into one or the other of two mutually exclusive classes. In
other words, every subject in both groups obtains one of two possible scores.
The scores are represented by frequencies in a 2 x 2 contingency table.'

Where the expert does not use the rule (as in some of the data) we have no
information. In the following example A, B, C and D stand for number of
posters.

\[
\begin{array}{c}
\text{Number of posters} \\
A & B & = & A+B \\
C & D & = & C+D \\
\end{array}
\]

\[
\text{Number of posters} \quad A+C \quad B+D \quad N \quad \text{Total number of posters}
\]

A is the number of 'not bad' posters computed by subtracting the number of
posters in which the rule actually was violated from the number of posters
to which the rule applies.

B is the number of posters in which the expert's rules were observed to be
violated in the rule based application.

C is the number of posters which were reported by the expert not to have bad
symptoms.

D is the number of posters reported by the expert to have bad symptoms.
**Analysis of the data using the Fisher test gives the following results:**

<table>
<thead>
<tr>
<th>Expert (A)'s rules</th>
<th>Expert (B)'s rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A1) No information</td>
<td>(B1) p = 0.61</td>
</tr>
<tr>
<td>(A2) No information</td>
<td>(B2) p = 0.0000003 (significant result)</td>
</tr>
<tr>
<td>(A3) p = 0.68</td>
<td>(B3) No information</td>
</tr>
<tr>
<td>(A4) No information</td>
<td>(B4) No information</td>
</tr>
<tr>
<td>(A5) No information</td>
<td>(B5) p = 0.37</td>
</tr>
<tr>
<td>(A6) p = 1.0</td>
<td>(B6) p = 0.002 (significant result)</td>
</tr>
<tr>
<td>(A7) p = 0.12</td>
<td>(B7) p = 0.45</td>
</tr>
<tr>
<td>(A8) p = 0.16</td>
<td>(B8) No information</td>
</tr>
<tr>
<td>(A9) No information</td>
<td>(B9) p = 0.47</td>
</tr>
<tr>
<td>(A10) No information</td>
<td>(B10) No information</td>
</tr>
<tr>
<td>(A11) p = 0.0002 (significant result)</td>
<td>(B11) No information</td>
</tr>
<tr>
<td>(A12) p = 0.08</td>
<td>(B12) p = 0.67</td>
</tr>
<tr>
<td>(A13) p = 0.009 (significant result)</td>
<td>(B13) No information</td>
</tr>
<tr>
<td>(A14) No information</td>
<td></td>
</tr>
<tr>
<td>(A15) No information</td>
<td></td>
</tr>
<tr>
<td>(A16) p = 0.13</td>
<td></td>
</tr>
<tr>
<td>(A17) p = 0.76</td>
<td></td>
</tr>
</tbody>
</table>

In the above a low p value is evidence of association. On this basis expert (A) uses 11 and 13 and expert (B) uses 2 and 6 consistently. In all other cases we have no reason to reject the null hypothesis that the designer does not use their rules to judge the designs.

This was an unexpected negative result which could have seriously undermined the thesis that elements of graphic design aesthetics are rule-based, and can be built into computer systems.
As will be seen, the follow up interviews produced an equally unexpected positive result. In these the researcher pointed out to the designers that their rules had been violated, and it turned out that the designers agreed 100% that their rules were being correctly applied.

7.5 Follow-up interviews

Analysis of the data revealed discrepancies perhaps due to ambiguities in some of the definitions derived from both experts' comments. To facilitate verification and clarification of these issues a follow-up interview with the experts was convened. The follow-up interviews benefited from the researcher's acquaintance with the experts and familiarity with the graphic design culture. On these occasions all the rules were discussed.

Follow-up interview with (A):

The elicited rules were discussed with (A) at his office in a meeting after the implementation tests were completed. The duration of the meeting was approximately 1 hour of which 30 minutes was spent in interrupted, informal conversation prior to focussing on the rules. (A) agreed to the tape recording of the proceedings.

The discussion revealed inconsistencies in the expert's reporting of bad symptoms and conditions which affected the application of rules and their interdependence. The rules which had been elicited were not reported to the expert before the follow-up interview; neither were the rules elicited from the other expert; these were noted but remain for future work., see chapter 9.
Further revision of the (A) rules:

After the follow-up interview, several of the rules were revised, adding conditions of their application. The definitions were also refined e.g.,

(A13) Original:
The words in a line should be separated by less than or equal width to a lowercase 'n' in the fount used; excepting lists and tables.

Revised:
The words in a line should be separated by less than or equal width to a lowercase 'i' in the fount used; excepting lists and tables.

Expert A acknowledged that there were anomalies in the reporting of some bad symptoms e.g. although it was agreed that

(A8)
Blocks of capital letters should be equal to or less than 4 lines.

This was reported on only one occasion when the expert gave the following reason for the bad symptom as

'The bottom 7 lines are unreadable'.

Although reported by Expert A to be important this symptom was not reported in three other instances where the researcher observed that the rule applied. In the systematic application of the rule these bad symptoms were routinely
identified, reported and shown as mismatches in the result of the implementation.

**Possible other factors in the identification of this bad symptom.**

i. The lines of type are centred; so are the three other instances in the other specimens.

ii. The type was set solid i.e. with no additional spacing between consecutive lines of type.

iii. The type size is 24 points.

iv. The block of capital letters reported is positioned at the bottom of the page.

v. The type style is Times, a serif typeface. Would there have been a difference if a sans serif face had been used?

The association of these factors and perhaps others which contribute to the symptom being reported as bad was not explicitly addressed at the follow-up interview. However it was observed that the other three instances had some leading (additional line spacing); it is evident that the addition of leading would improve the readability of this block of type.

However a block of three lines of type in the same layout and with the same type parameter as above was not judged by the expert to be bad, so the number of lines (between 4 and 7) is the crucial factor in this instance.

Enhancement of the rules and their re-application to the sample of posters produced complete matches in the observed instances.
Follow-up interview with (B):

The same procedure was employed as applied to the meeting with (A). The duration of the meeting was approximately 35 minutes. (B) agreed to the tape recording of the proceedings.

The researcher showed the list of rules to (B) and both referred to them during the discussion. The discussion with (B) revealed conditions which affected the application of rules and their interdependence as had the interview with (A).

The rules elicited from expert (A) were not reported to expert (B).

The new information obtained at the follow-up interview was used to enhance the formulation and subsequent application of the (B) rules. There were two notable instances of rule enhancement.

(B13)
Blocks should be parallel to the horizontal edges of the page or at an angle greater than 10 degrees from the horizontal.

was modified to:
The horizontal edges of blocks should be parallel to the horizontal edges of the page.

(B2)
If there are top and bottom margins, then the top margin should be less than or
equal to the bottom margin.

was modified to:

*If there are top and bottom margins, then the top margin should be less than the bottom margin.*

(B7)

If there are three or more vertically adjacent blocks of the same fount then the distance between the adjacent edges of any two of these blocks should be less than $2.1/2 \times$ the type size in these blocks.

was re-formulated to:

*The vertical space between three or more adjacent blocks of the same type size should be less than $2.1/2 \times$ the size of type used in these blocks.*

But the modification devised remained ambiguous and was eliminated from the list for the following reasons.

i. One problem was that it was difficult to determine accurately the type size without reference to a type specimen sheet or catalogue.

ii. Another problem was to determine whether the vertical distance (of $2.1/2 \times$ type size) should be measured between the base line and the mean line of different adjacent pairs of lines of type or between the base of a descender and top of an ascender.

iii. Yet another problem was in the application of this rule to blocks of capital letters which have neither mean line nor ascenders and descenders.
However, enhancement of the rules produced complete matches in the observed instances in the re-application of the re-formulated rules to the sample.

7.5.1 Conclusions of the follow-up interviews:

The results show that for the two designers studied,

(C1) These designer's intuitive judgement of bad design in a population of A4 desktop publications can be elicited and formulated into a coherent and rigorous set of rules which are generalizable to disparate layouts within the population.

Analysis of the results shows that:

(C2) These designers (perhaps unknowingly) use rules to diagnose bad design symptoms.

(C3) These designers' rudimentary intuitive knowledge can be interpreted as rules.

(C4) From their comments, the experts attempt mentally to recompose a layout when assessing its defects before offering their prognosis. Indeed alternative solutions were sometimes offered. Also the experts evidently attempt to determine the prevailing constraints in an attempt to evolve a more appropriate alternative solution. This can be deduced from the recommendations and suggestions they offer in their reasons why a symptom is bad.
A graphic designer consistently reuses some of the same ‘rules’ in the evaluation of different specimens.

A graphic designer’s perception of layout design is conditioned by the number of bad symptoms present in a layout, i.e. the fewer the number of defects the more forgiving the assessor. Consequently a major defect is sometimes overlooked where there are few or no other bad symptoms. Where there are many bad symptoms the experts diagnosis is more stringent.

Graphic designers apply their rules inconsistently to different specimens; the computer rule based system is more reliable.

A rigorous method has been described from which it can be concluded that few of the rules tested are predictable. However, the evidence suggests that the performance of the rules benefits from feedback to their source i.e. it was observed that rule application was highly consistent with the expert after the follow-up interview.

It has been demonstrated that the systematic procedure is more consistent than the expert in the application of rules formulated from the expert’s comments.

However expert (A) acknowledged that he was not always aware of or able to articulate the subtlety of his conditions for the diagnosis of bad symptoms.
This it is suggested contributes to the expert's indecision and results in uncertainty in his diagnosis of bad symptoms.

*Predicting the judgements of the experts (before the follow-up interview).*

We can add all the “bad” predictions to get a single number, and see how well we predicted violations being judged:

Number of bad symptoms predicted by rule-base (A) = 104  
Number of bad symptoms reported by expert (A)  
(and included in the rule-base) = 32  
Agreement = 30.76%

Number of bad symptoms predicted by rule-base (B) = 87  
Number of bad symptoms reported by expert (B)  
(and included in the rule-base) = 17  
Agreement = 19.5%

*Results adjusted after verification with experts in the follow-up interview:*

During the follow-up interview the application of the rules formulated from the experts comments was verified. Where the application of some rules was previously uncertain their application was confirmed and applied more consistently in a subsequent re-evaluation of the sample.

This resulted in an increase in the number of bad symptoms observed in the
rule-based re-evaluation using expert A’s rules and a decrease in the re-evaluation using B’s rules.

Number of bad symptoms observed by rule-base (A) = 119
Number of bad symptoms agreed by expert (A) at the follow-up interview = 119
Agreement = 100%

Number of bad symptoms observed by rule-base (B) = 79
Number of bad symptoms agreed by expert (B) at the follow-up interview = 79
Agreement = 100%

One interpretation is that the prediction is poor, until the expert is asked to consider the particular rules. An alternative view is that the rule-base is more consistent than the expert in the application of their rules.

7.5.2 A follow-up test to see if expert A uses his rules consistently after ten months

Are experts consistent in their re-diagnosis of bad symptoms in a sample?

To examine this issue a test was devised to determine whether the experts would diagnose the same symptoms in a re-evaluation of the sample. It was decided to verify this hypothesis by requesting one of the experts (expert (A)) to repeat the evaluation. Expert (A) was available and agreed to re-evaluate
the sample. By this time ten months had elapsed since the previous test so it was thought unlikely that the expert would remember his previous responses. Also it had been confirmed that the expert had not copied their previous comments; the effect of which would invalidate the new test. The results are shown in table 7.1 page 176.

After the follow-up interview when the rules were endorsed, it was expected there would be an overall improvement in the expert’s performance in the re-evaluation of the sample i.e. the expert would diagnose most or all the bad symptoms after a teach-back ‘training’.

Comparison of the data suggests that:

a. the expert’s performance in diagnosing some bad symptoms is inconsistent and therefore unpredictable.

b. some rules gain in importance at a subsequent application.

c. the expert does not diagnose all the bad symptoms.

The expert’s evaluation only captures some of the bad symptoms.

Moreover the expert does not consistently re-apply the same critique to previously examined specimens.

However some of the bad symptoms which were not identified in the first evaluation of a sample were found in a second iteration. Additional bad
symptoms can be found by adding the bad symptoms in two or more evaluations of a sample in a population of designs. Evidently by repeatedly testing the expert more bad symptoms may be revealed.

Where a second test is performed the expert's diagnosis may not agree with the results obtained in the first test. This suggests there is a time related nature of evaluation which perhaps has consequences in other domains.

An alternative view is that data from a repeated evaluation is not additive i.e. the expert's diagnosis of bad symptoms in one sample cannot inform their re-evaluation of the same sample or the diagnosis of bad symptoms in a new sample.

If this view is accepted, then bad symptoms elicited from different samples or from the sample on different occasions are analogous to unique snapshots of the expert's preferences where each is taken from a different viewpoint and perspective and therefore cannot be assembled to form the coherent universe of the experts rules.

However, while the expert's values at any given time may be fixed, over time it has been shown that it is susceptible to change.

Change in the addition of new rules or removal of others informs the enhancement of the expert's rule set and its application.
7.5.3 Some possible causes of the experts' inconsistency

What were the reasons for the experts' inconsistency in the application of their rules? The experiment shows that the expert's judgement is not absolute, it is provisional and inconsistent.

Inconsistency suggests that the expert is perhaps uncertain of the application of the rule and/or the factors affecting its application. It may also reveal the addition of new rules. But this is difficult to verify. It may show the withdrawal of a previously used rule. The addition or withdrawal of a rule in regrading marks the point when, where and in what circumstances a rule is used.

The experts agreed with the rules formulated from their critiques of the specimens in the samples when their rules were discussed at the follow-up interview.

This was perhaps due to higher standard of design reported by the experts. As reported earlier, the result was that the experts were more forgiving of the bad symptoms observed overall in the second sample.

In some instances the experts had not applied their 'rules' because there were additional factors which conditioned the rule's application in the second sample.
There is evidently inconsistency in the experts' application of their rules resulting from their difficulty in determining the equality of margins.

The experts evidently varied the weighting of the rules according to the importance they attributed to the defect and its codependents in different layouts.

Inconsistency is perhaps due in part to the monotony of the repetitious nature of following a diagnostic procedure. The variability of the sample may be another factor e.g. some contained imagery while others were text only.

Inconsistency may result from visual illusions and perceptual constancy. The expert's cognitive style may be field-dependent or field independent i.e. in the perception of a publication some people may be more affected by the content and others by the page. Inconsistency may be due to interest in or dislike for the subject of the narrative. The expert's attitude may be affected by low motivation resulting from the lack of interest and lack of reward.

However in this study we must assume the expert's integrity. There were also evidently differences in the expert's ability to recognise the difference between the sizes of top and bottom margins i.e. expert (B) reported 8 defects while expert (A) reported none, although both reported this defect in the first sample. This would be a relatively rudimentary issue for a systematic procedure to address.
7.5.4 Summary

This study describes a novel method for the elicitation of aesthetic rules of graphic design evaluation from two graphic design experts. The procedure devised relies on the graphic design experts' independent diagnosis of symptoms of bad design in a sample of graphic layouts.

The diagnosis of 'bad' design symptoms in a graphic layout is shown to be an effective method of making explicit the graphic designers' implicit intuitive, aesthetic judgement explicit. It is argued that the presence of a bad design symptom in a graphic layout indicates that the design is not good and that the absence of bad design symptoms suggests that the design is not bad.

Pye (1978) reports that 'people do not unanimously agree about what is beautiful and what is not, for they do not unanimously agree about anything whatever'. Here it has been demonstrated that complete agreement between the successive critiques of the same expert is unlikely and suggests a reason that agreement in different individuals' opinions should not be expected. This suggests why a rule base performs aesthetic evaluation with greater consistency and reliability than experts.

The procedure devised required the graphic design experts to provide their rationale for their diagnosis and this was used to inform the application of their rules to the bad symptoms identified in the graphic layouts.
The subtle expressions used by the graphic design experts provided interpretational difficulties which were resolved by the formulation of the rules into representational statements suitable for use in a computer system using the methodology of 'The Expert System Test'. This methodology supports the unambiguous application of a rule in the diagnosis of a bad symptom.

Recent research supports the viability of automatic procedures for document-layout analysis.

The design and implementation of rigorous statistical procedures suggests that the graphic designer's aesthetic values can be systematized.

The result of tests show that the systematic application of graphic design rules performs more consistently in design evaluation than the experts from whom they were elicited.

From the above data it can be observed that:

- Not all rules can be applied to all posters.
- Rules can be elicited from an expert and consistently applied in a rule based system.
- Designers do not apply their rules consistently.
- When a designer has the rules made explicit, he or she will apply them in 100% agreement with the rule based system.
- When one designer was asked to evaluate the rules after 10 months, he reverted to great deviation from his set of rules.
When the design rules are implicit the designer applies them inconsistently. When the rules are made explicit the designer applies them consistently.

The rules are acceptable to a graphic designer representing “canons” which they believe, i.e. they are real aesthetic (bad, not bad) design rules. Because they are explicit rules, they can be built into a computerised graphics system to support more naive users.

7.6 Conclusions and discussion

Difficulties revealed by the experiments

Importantly, rules A1 and B1 relating to the inequality of left and right margins was found to be wrongly included in the rules because the experts had not reported it with the statistical frequency required. Further analysis of the data revealed that on one occasion the experts remarked on a defect which resulted in the formulation of this rule. However, when the rule proved to be unreliable its inclusion was not challenged by the experts in a follow-up interview. Indeed, during the discussions with the experts, each confirmed that they were concerned about the inequality of left and right margins but could not explain why they had not diagnosed this bad symptom.

The results tell us that some intuitive design judgement is variably applied:

i. The experts failed to report some bad symptoms consistently.

ii. Some rules were applied inconsistently.
iii. Several new symptoms were reported in the implementation which were not mentioned in the test.

iv. The interdependent conditions associated with some rules and the criteria surrounding their application were uncertain/indeterminate from the experts’ written comments.

**Shortcomings of the rules-based evaluation revealed at the follow-up interview:**

Two other problems were that:

i. Four of the definitions were ambiguous and required enhancement.

ii. Six of the rule statements required simplification.

These issues were addressed in the follow up interview when it was possible for the researcher to tease out from the experts some of the conditions affecting the application of the rules.

The problem is that some formalised rules are initially unrecognised when fed back to designers from whom they were elicited. This is due in part to the need to separate the expert’s comment into two or more components to form logical expressions which could be unambiguously applied.

The measurements and tolerance applied to the features in a layout e.g. margins, were manually measured by the researcher but were reported by the
experts to have been visually compared to determine congruency and alignment. The experts appear to apply these rules inconsistently due to uncertainty in distance measurements. Also because the layout is sometimes askew on the page it was often difficult for the researcher (and evidently the expert) to determine the appropriate point(s) at which to measure margin conformity; particularly where the space along a side margin is just outside the 2mm tolerance specified. Perhaps the 2mm tolerance is too small to be visually detected in the various combinatorial configurations presented in an A4 publication.

Alternatively, it is possible that the experts may have decided through their personal experience of the limitation of the process of reproduction, that equal left and right margins could not be expected from a typical photocopy. However expert (B) was insistent on the top and bottom margin ratios being adhered to.

Photocopies:
The process of photocopying proved to be the most economic method of duplicating the sets of specimens required. It was relatively quick and economic to produce copies which looked dimensionally similar to the originals. The viability of this approach had been tested in the pilot test. Before the implementation the approach was reviewed. It had been noted that photocopies were not exact replicas. Some machines produced copies which were smaller while copies from another would be larger yet others would be
distorted and others would produce copies with all three defects. Another problem was that originals placed on the same relative position on different machines produced copies which had different margin widths. A faithful reproduction, dimensionally exact was impossible to achieve using this approach.

Clearly this approach is not satisfactory. On further consideration the specimens could have been photocopied more carefully; perhaps onto A3 and cut out to size.

An alternative would be to make printed copies from the original posters. This would minimise the variations in position of the image elements on the duplicates. It would also eliminate any distortion in the shape of the composition on the page. However this process requires specialist expertise and is costly in resources and time.

**Constraints on expert's comments:**

The size of the questionnaire restricted the expert's response; their comments were synoptic, where a more comprehensive report would have provided information concerning the influences and conditional issues which were mainly ignored. However the follow-up interview did reveal additional information which helped to clarify some of the more difficult points concerning the precision of definitions and emphasis within a rule. The experts were required to provide written expression for their attitudes and
beliefs, some of which were perhaps transient and vague. They had to convert their emotive feelings into reason. Many design values are based on convention and tradition. However, some design traditions are difficult to rationalize e.g., that centred type should be used only for the poetic, the romantic, for formal occasions and subjects with classical pretension.

It is evident from the data available that some of the results obtained were statistically inconclusive e.g., there were no bad symptoms present on some specimens. On other specimens, the number of bad symptoms diagnosed by the expert and or the rule-base gave values which were too small to provide conclusive results.

The experts' incomplete explanation of the symptoms and the codependent issues affecting the application of the rule impeded the formulation of comprehensive rule statements.

The results underline the importance of the follow-up interview to the verification of the rules formulated from the experts' comments after which consistent results were obtained by the re-application of the enhanced rules. This suggests that if the experts were themselves requested to apply a check-list of rules formulated from their reported comments then the result would be in agreement with that obtained by the rule-based application.

It is evident that the expert does not use an explicit list of rules. He or she
responds to observed cues in the layout which prompt a reaction to examine the suitability of elements represented against specific values and precepts.

Identification of the bad symptom is dependent on its visibility. There may be a variety of bad symptoms. Symptoms present differently. Bad symptoms may be obscured in a complex composition and highlighted in a spacious deployment of elements. The close proximity of elements is an aid to determining vertical and horizontal alignment. The further apart elements are situated the more difficult it is for the expert to detect configurational deviation with certainty. This may result in a defect being overlooked or forgiven where there is uncertainty. Where for example there are several lines of centred type in a block it may be more difficult for the expert to determine whether left and right edges are the same distance from the edge of the page than when the text is justified.

Recognition of a defect may be affected by surrounding elements. Horizontal misalignment may be obscured by the presence of an unsuitable word spacing in a text block for example.

The rule-base on the other hand, actively interrogates all the elements when searching for bad symptoms. It can search through the document and page set-up data, element by element to extract the number of typefaces from the list of specified type parameters. Similarly the number of type sizes can be accurately determined.
Nevertheless, it could be argued that systematisation is limited by its reliance on the inadequate interpretation, distillation and simplification of complex, codependent, intuitive, variable design expressions and incomplete comments which have been shown to be difficult to translate and make objectively explicit. Until these problems are resolved a systematic approach will not attain the academic credibility and professional currency necessary to qualify as a reliable recipe for the declaration of good design.

However, despite these issues it has been demonstrated that a systematic diagnostic procedure can accurately inform the user of the presence of prescribed bad symptoms in a publication. Indeed it is argued that a high (consistent) performance is offered by the systematic procedure which may be preferable to the low (inconsistent) performance attainable by the expert in design evaluation for the reasons given.

The expert's diagnostic procedure is unreliable and ineffective because it fails to identify all bad symptoms. It is evident that bad symptoms are sometimes missed in the apparently random (disorganised) search of a layout in the expectation of witnessing bad symptoms. Moreover, the expert's inconsistency in reporting bad symptoms suggests that in the search for defects, the content and detail in a layout are examined arbitrarily i.e. the expert opportunistically searches for a variety of conditions which might present bad symptoms in a layout. However although the rules are not arbitrarily applied the experts report that the evaluation process and reporting of bad symptoms
was itself inconsistent due to fluctuating mood and attitude to the layouts. The experts’ empathy with the content is also reported to have affected their responses.

The experts’ critique is however an effective process for the exposition of physical and emotive beliefs through the diagnosis and reporting of bad symptoms in a sample of publications. Much of the experts’ judgement revealed in their comments can be translated and employed in a reliable systematic diagnostic procedure.

The diagnostic approach evidently reveals the presence and absence of rules through their application or non-application, agreement or violation. The results obtained inform the systematic validation of layouts based on the intuitive judgement of experienced designers.

The experimental methods described enable the researcher to determine the similarities, differences and effectiveness of the diagnostic rules elicited from two or more sources.

The experts’ diagnoses of bad symptoms suggest that they have a preconception or mental model of good layouts (and good design) perhaps based on their extensive graphic design experience. They may also have the ability to mentally compose a better alternative solution than represented in the specimens examined. Therefore the diagnoses of bad symptoms suggest a
mismatch between element(s) in the layout and the experts' expectation or prescription.

It is concluded that apparently trivial rules are consistently obeyed by this group of designers. Furthermore, it is evident that trivial rules are applied more consistently than subtle rules.

Dreyfus and Dreyfus (1986) have asked the question 'Do expert systems ever do as well as the experts whose rules they run?' The reply in this instance is affirmative. The results show that design rules can be elicited from graphic design experts and have been shown to perform more consistently in design evaluation than experts from whom they were elicited; by consent of the experts themselves.

Finally, it has been demonstrated that

1. Designers judge bad results according to their rules.
2. Designers don't apply their rules at random.
3. Designers agree with the prediction of (some) of their rules.
4. Some rules are not very good at predicting because:
   a. They did not capture the designer's meaning adequately.
   b. The weighting on the badness was inconsistent.
   c. The "badness" seems to be additive over a number of violated rules.
d. Special external conditions may “excuse” the rule.

e. The expert reported bad symptoms inconsistently.

5. Simple rules could be added to DTP systems and would help to avoid making elementary design errors. The computer would apply these more systematically than the human designer.

6. In this investigation systematisation has been observed to have the potential to:

a. Ensure that prescribed values are routinely applied. This would contribute to the elimination of confusion arising from inconsistent human sources.

b. Provide impartiality - it is uninfluenced by sentiment; visual or narrative appeal. It is not susceptible to empathy or novelty - applies the same conditions to old and new styles; the familiar and obscure are treated with equal impartiality.

c. Reflect and support personal or collective views.

d. Be explicit and can provide rational advice.

e. Be routinely efficient - permits speedy identification of contraventions.

f. Be uninhibited and even-handed - declares the bad symptoms to all users with the same conviction.

7. Expert evaluation has been shown to be:

a. Arbitrary and conditional - as observed in the changing importance assigned to the symptoms reported by the experts.

b. Variable and incomplete - to the extent that the obvious is often overlooked.

c. Biased - as shown in the differences in the two sets of rules elicited from the experts.
d. Discretionary and forgiving - the reporting of defects is conditioned by the nature and the number of observed incidents. This was confirmed by the experts at the follow-up interview who report being influenced in their responses by their perception that the second set of specimens were 'much better designs than the first' and were more tolerant to bad symptoms i.e. where there are few bad symptoms the expert may ignore the defects.

e. Superficial - design evaluation is seen to be fleeting, cursory and dismissive. The expert appears to this observer, to be making it up as he goes along because of the variable time reportedly spent in the evaluation.

However, the expert may be dismissive when there are many bad symptoms present in a specimen.
CHAPTER 8: CONCLUSIONS

8.1 Introduction

This study describes a novel method for the elicitation of aesthetic rules of graphic design evaluation from graphic design experts. The procedure devised relies on the graphic design experts’ independent diagnosis of symptoms of bad design in a sample of graphic layouts.

It has been shown that the rudimentary implicit, intuitive, idiosyncratic and apparently arbitrary rules used by graphic designers for the aesthetic evaluation of desktop publications in the A4 population of posters and notices can be elicited and systematized. However it has also been demonstrated that few of the graphic design expert’s rules are applied consistently, unless the flaw is pointed out in which case they are consistent.

8.2 Summary of research findings

This thesis makes contributions in the following areas:

- It has been shown that:
  a. The diagnosis of bad symptoms in a graphic layout is an effective method of making explicit part of the graphic designer’s implicit aesthetic judgement.
  
  b. Some of the natural language and specialist expressions obtained in reports of bad symptoms can be formulated into logical rules suitable for systematic application.
  
  c. Graphic design experts use a form of intuitive rules to perform aesthetic evaluation.
  
  d. A design cannot be definitively classified as good but can usefully be classified as ‘not bad’ with respect to a given set of rules each of which can be applied independently.
• Graphic design experts are inconsistent in the application of their rules over periods of time (ten months).

• Graphic design experts alter their perception of design defects over time. In the re-evaluation of a sample of specimens the expert’s values unknowingly changed in stringency.

• At any instant of time different graphic design experts have different perceptions of some design symptoms and similar views of others, i.e. experts share some aesthetic rules.

• Graphic design experts do not diagnose all the bad symptoms in a graphic layout evaluation task.

It is concluded that:

1. Graphic designers employ intuitive, implicit aesthetic rules in the evaluation of a graphic layout which can be made explicit.

2. Graphic design rules can be applied systematically to a population of graphic layouts.

3. Expert graphic designers have rules, but they apply them inconsistently.

4. Graphic design aesthetic expertise can be elicited and entered into a computer system to enact the role of a graphic expert.
CHAPTER 9: FURTHER WORK

9.1 Introduction

The methodology described informs the systematisation of the interdependent rules and the simple application of 'bad symptom rules'. The following describes ways in which this prototype system could be further developed.

There are several commercially available visual design DTP systems to enable novices and experts to prepare a variety of desktop publications. But these systems presently provide no design evaluation support.

Recognising this deficiency Monotype, (see Waller (1991) a major manufacturer of typesetting equipment and software, have produced a package together with supporting documentation which offer some graphic design rules and rudimentary guidance to support novices in the preparation of desktop publications. Systematic design evaluation support would be beneficial to this class of desktop publishing users.

In a prototype system, a human application of the rules simulating a simple automatic graphic design evaluation procedure has been demonstrated. But clearly, further enhancement of the rules and the procedure is possible and necessary.
9.2 Further research

This study relied on the diagnosis and comments on bad symptoms by two experienced graphic designers. Both of the expert’s responses contained comments and statements of the importance of the various ‘bad symptoms’ diagnosed. From the analysis and computation of this data the rules used may be organized in a hierarchy to represent their relative importance as indicators of bad design. The complement of rules could be subdivided into logical sets to address heterogeneous properties e.g. alignment and margins. Groups of rules could be arranged in interdependent clusters e.g. text elements. Weighting could be assigned to the rules according to the user’s perception of their importance in the evaluation. Here the list of rules could be presented for the user’s grading.

The procedure could be designed to enable key areas to be addressed selectively. The user could elect to concentrate on the characteristics of a specific block of text to speed the consultation. The diagnostic procedure could provide prognosis after a full document analysis rather than giving an interim report when a bad symptom is diagnosed. Indeed users may prefer to be informed about all the bad symptoms collectively at the conclusion of a diagnosis.

While the multi-purpose Leonardo expert system shell used for the partial simulation of the evaluation procedure was adequate for the partial prototype simulation of the rule-based system it was inadequate for the graphic visualization or identification of features in a layout. Therefore a more
sophisticated system incorporating image recognition, layout analysis, and graphic representation would be essential.

An important characteristic of expert systems is their ability to engage in a coherent, plausible, reciprocal exchange with the user simulating rudimentary human dialogue. This feature provides the potential for use at two levels; one level is that of a tutor to novices; students and non-designers. At a higher level, it can be developed into a robust interactive design assistant which not only volunteers advice but which can discuss alternatives through query/response algorithms or routinely remove bad design systems as the design procedure progresses.

As an intelligent tutoring system, a design evaluation system could advise the user of bad design symptoms with unerring consistency. The availability of a rigorous method of graphic design evaluation informs the design of a sophisticated automatic procedure to support design evaluation at a variety of levels.

Non-designers and students of graphic design are often unaware of the fundamental rules on which successful graphic layouts depend. For these groups the preferred visual configuration could be presented in addition to the appropriate objective advice being given, making explicit traditionally unstated perhaps universal beliefs, and thereby accelerating the learning process.
A design evaluation procedure may be independent or integral to a desktop publishing package. An independent system could be designed to interrogate prescribed elements of the completed layout while an integral system would examine each detail as it was created; by systematically identifying and notifying rule violations.

There are advantages and disadvantages associated with both procedures. An independent evaluation system could be seen as being on-line but operating as a call up facility of an existing desktop publishing package. An integral evaluation system would be an integrated application of a desktop publishing package. This can be likened to a Dolby system in a stereo amplifier which can be activated or deactivated at the listeners discretion. While the former would offer flexibility, compatibility could not be assured. So-called compatibility is notoriously difficult to guarantee. The integral system would guarantee compatibility of the whole system.

The expert system may be designed to be customized by the user, incorporating the user’s preferences or those of a chosen design expert. The expertise contained in the expert system could be permanently or temporarily altered by the user. The effect of any alteration to margin ratios, leading and readability parameters for example, could be computed automatically by the system and presented as a visual display message, information or query.
A robust evaluation system could be used to compare one designer's preferences to another. It also enables the user to compare the choices made by two or more graphic designers and the consequences of their choices on the legibility and readability for example, on the users layout(s).

This form of analysis enables the user to examine the result of changing any one, or a number of the graphic design decisions. If we know enough about the properties of the design elements and the nature of the narrative, in terms of its formality or informality, we can more reliably determine the designers intuitive reasoning and better understand the nature and number of decisions the graphic designer has to make.

By making two or more evaluations for example, based on the rules of different designers the user could be provided with a unanimous or majority judgement on a layout's merits. However, the user may be surprised by the inconsistency and variability of human evaluation.

The principal difference between a human evaluation of layouts and evaluation by expert systems is that while humans are fallible, variable, and unreliable, an expert system can be programmed to remain consistent in the evaluation of a given population of layouts. Automatic processes remain unimpressed by first impressions, the presence of grammatical errors, image quality and incoherent narrative. Humans are apparently preoccupied by these issues which contribute to variations in their judgement. Human variability is vital for
innovation and the generation of new initiatives and offers the opportunity to take advantage of chance occurrences. The consistency offered by automatic evaluation systems is essential to ensure that typographic details in a layout for example, are not overlooked and that a predetermined aesthetic quality is maintained.

In the expectation that systems of automatic page layout will be evolved it is envisaged that an integrated sophisticated, robust, systematic procedure for aesthetic evaluation could provide routine design support particularly for visual design systems.

This complementary automatic aid to graphic design evaluation remains for further research.
REFERENCES


228


229
FIGURES AND TABLES

Figure 2.1: Halprin's model of the environmental design process 23
Figure 2.2: March's model of the architectural design process 25
Figure 2.3: French's model of the mechanical engineering design process 27
Figure 2.4: Pahl and Beitz's model of a traditional graphic design process 29
Figure 2.5: Glaze's model of a traditional graphic design process 46
Figure 2.6: Model of Tunnicliffe's computer assisted graphic design process 51
Figure 2.7: Model of Cotton's manual and computer aided graphic design process 52
Figure 3.1: The Golden Mean 66
Figure 3.2: The Golden rectangle 66
Figure 3.3: The Golden Rule 1: 1.618 is applied in this example to the ideal proportions of a page 68
Figure 3.4: The term Golden Section when applied to the pages of a book 69
Figure 3.5: Garlands type constraints on an A4 page (Garland 1969) 71
Figure 3.6: Ed Day's grid designed for a Mitchell Beazley encyclopaedia 72
Figure 4.1: Hopgood's (1993) model of a rule-based system 98
Figure 4.2: Forward chaining 102
Figure 4.3: Backward chaining 104
Figure 6.1: Six of the forty A4 poster specimens 133
Figure 6.2: Specimen layout showing abstract constituents of text and picture elements 135
Figure 7.1: The reiterating process of rule formulation 164

Table 1.1: Rules and options in major DTP and word processing software 9
Table 6.1: Rubinstein's recommendations violated magazines and commercial produced leaflet samples 131
Table 6.2: Rubinstein's recommendations violated in forty A4 specimens 134

Table 6.3: Comparison of two subjects evaluation of the DTP publications 143

Table 7.1: Comparison of a rule based diagnosis using (A)'s rules, with his repeated diagnosis of bad symptoms in the same sample 176

Table 7.2: Comparison of a rule based diagnosis using (B)' rules, with her repeated diagnosis of bad symptoms in the same sample 177
This thesis is produced using a Macintosh LC475 computer. The text copy is set in 12point Times u&l.c. justified; bold is used for headings, bold italic for subheadings and italic is used for emphasis. The chapter headings are set in 18point Palatino caps and the initial letter in 36point Times. The software used is Quark Xpress 3.11. It was printed on a 300 DPI Macintosh Laserwriter II NTX. The page layout conforms to the requirements of the Open University Higher Degree regulations: margins and line spacing, etc.