ED 238 686 SE 043 549

AUTHOR	Simcox, William A.
TITLE	A Comprehensive Process for Display Systems Development.
INSTITUTION	Consulting Statisticians, Inc., Wellesley, MA.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Jan 83
CONTRACT	400-79-0066
NOTE	75p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Communications; *Computer Graphics; *Design Requirements; Display Aids; *Display Systems; *Human Factors Engineering; Visual Learning; *Visual Literacy; Visual Perception

ABSTRACT

A comprehensive development process for display design, focusing on computer-generated cathode ray tube (CRT) displays is presented. A framework is created for breaking the display into its component parts, used to guide the design process. The objective is to design or select the most cost effective graphics solution (hardware and software) to meet any problem situation. The framework is divided into informational structure/medium of display; these are broken into constituent elements and discussed with associated design issues. The informational structure includes content, format, and organization. The CRT is analyzed for display design criteria such as light, geometric parameters, and CRT display properties. Viewing area criteria are discussed. The components of medium and informational structure provide the framework for specifying hardware/software requirements for graphics systems. The process of display design is presented, beginning with feasibility studies, resulting in a set of useful solutions. Steps to obtaining a 'best' solution are summarized, creating an optimal candidate system which is physically realized and tested through detailed design activity. Results may indicate a need for revision/redesign requiring alternative solutions. The criterion for selecting the most effective solution is based on integrity of design. A list of 24 references, 18 figures and 6 tables are attached. (JM)

A COMPREHENSIVE PROCESS FOR DISPLAY SYSTEMS DEVELOPMENT

NIE 400-79-0066

Year 03

January 5, 1983

Submitted to:

The National Institute of Education
Contracts and Grants Management Division
1200 19th Street, N.W., Room 808
Washington, D.C. 20208

ATTN: Martin Engel, Ph.D.

Submitted by:

William A. Simcox
Senior Engineer
Consulting Statisticians, Inc.
20 William Street
Wellesley Hills, MA 02181

A COMPREHENSIVE PROCESS FOR DISPLAY SYSTEMS DEVELOPMENT

INTRODUCTION

In this, the age of the computer, the computer generated display is fast becoming the preferred mode of communication in many areas. At first these displays were found only in high technology environments such as aircraft cockpits, C^3 systems and the like. Gradually they found their way into the process control industry, both nuclear and conventional. Now we are even seeing computer displays in the educational setting as with computer aided instruction, and as a general information tool through the process of videotex.

This explosion in available display technology brings with it a host of decisions on design related issues. For example, which type of symbol generation technology should be chosen? Which display configuration will best communicate the information? What size display will be used? Before any display system becomes operational in an on-line sense, these as well as a host of other design issues must be resolved to effect the most cost effective design. The identification and resolution of these issues must not be based on intuition or brute force trial and error, but rather on a systematic development process which draws upon available facts and empirical data relevant to each decision step. Unfortunately, a comprehensive development process has been lacking in the area of display design, although initial attempts at developing such methodologies can be found (Hitt, Schutz, Christner, Ray & Coffey, 1961; Rogers, 1981). However, these attempts do not encompass all the relevant design issues.

The purpose of this paper is to present a comprehensive development process for display design. The comprehensiveness of this process is achieved by first introducing a framework that partitions the display problem into its constituent elements and then using this framework to guide the steps of the design process originally formulated by Morris Asimow (1962). The result of this approach will be the capability to design or select the most cost effective graphics solution--hardware and software--to satisfy the needs of any problem situation.

FRAMEWORK

The characterization of the display design problem begins with a delineation of the two components of a displa[illegible]t work together to communicate information. Proper execution of both components is necessary for successful information communication to occur. These components are called the informational structure and the medium. The informational structure refers to the representation that is actually being presented by the display. For example, an informational structure could consist of a graph of gold prices over time, a map of the United States showing average rainfall per region and the like. The medium refers to the device used to present the information. It makes the informational structure accessible. Most generally, the medium is a Cathode-Ray-Tube (CRT), but it can also be a page of paper, or poster board for example. Note the the medium, as stated before, affects the access of the information while the informational structure affects interpretation Thus it is clear why proper design of both is necessarily essential to effective communication and why design emphasizing either one alone is not sufficient.

In the remainder of this section, design relevant properties of each component are presented and their effects on access and interpretation are discussed. Following this, a general process for designing display systems is presented. This process, together with the framework, should allow one to design or select the most cost effective display relevant to a specified task or tasks.

Informational Structure

The informational structure is the information conveying representation of the display system. How it is designed bears an important relationship to ease of information interpretation. All too often, a poorly designed informational structure leads to an increased level of cognitive effort required by the reader to process the ambiguous information it contains. This increased effort leads to greater incidence of fatigue, which in certain cases can have catas-pnic consequences. Thus it behooves us to consider what properties of the informational structure make it more or less ambiguous and what steps can be taken to correct any deficiencies.

Most of the background for this section comes from our book Understanding Charts and Graphs: A Project in Applied Cognitive Science, specifically Chapter 2, written by Stephen Kosslyn. Concepts from that chapter will be used where appropriate in the remainder of this discussion.

Any informational structure can be specified in terms of three components --content, format and organization. The content of the informational structure is the underlying information to be conveyed. The format determines the manner or mode of conveying the information content, while the organization specifies the relationship among the elements of the format. Each of these components will be expanded upon in turn.

Content

Following McCormick (1976), some of the major types of information presented by displays are summarized below.

Quantitative information. Information which reflects the quantitative value or values of one variable over the quntitative value or values of another variable. We see this information in the form of common graphs and tables.

Qualitative information. Information which reflects gross changes in a quantitative variable. For example, direction of change or rate of change coded verbally would represent qualitative information. Note that while the underlying variable is quantitative, the information is used more as an indicator of change rather than in a precise way.

Additionally, qualitative information would take the form of structural/organizational representations such as those found in flowcharts and the like. Here the direction of and number of mappings are information conveying.

Status information. Information which reflects the condition or status of a system, such as: on-off indications, or indications of one of [illegible]ber of conditions of information. Warning information would be included in this category.

Representational information. Information found in depiction of objects, areas or other configurations. Any representation that conforms to Kosslyn's (1980) five criteria for depictions would be classed in this category.

Alphanumeric information. Verbal and numerical information such as instructions, tables or any printed and typed information.

The content of any display must be clear from the outset since certain formats are better suited to certain types of information than others. In some cases, the content constrains the kinds of formats that can be used, but in most cases, we have some latitude in our choice. What we mean by format is described next.

Format

The format of an informational structure refers to the form of the representation that makes the content of the information available. Following our terminology, the format of a display is determined by its framework, specfier and labels.

The framework provides the visual bases for the specification of relationships between the variables comprising the content of the information. In line graphs, for example, the framework is a pair of axes showing the entities being related and how they are measured. In charts, the framework is the different entities often represented symbolically by different shapes. For alphanumeric representation, the framework is the different alpha and verbal labels that make up the to be specified text. The specifier serves to bring the parts of the framework together in relationship. In a line graph, the specifier serves to pair a value or range of values on one mathematical scale with a value or range of values on another mathematical scale. The specifier in a chart would consist of the presence or absence of arrows connecting the various shaped boxes as well as their directions. Finally, labels serve to inform the reader of the actual entities being represented.

The design question concerns the appropriateness of the format chosen to the type of information conveyed. That is, for a given piece of information, how should it be represented in the display. Simcox (1982) has shown that an incongruency between format and response basis leads to an increase demand on our information processing resources and an increased incidence of classification errors. Kosslyn and James (1982) have obtained similar results using more realistic graphic displays. At this point, the choice of format must be empirical. However, with regard to quantitative information, Chapter 7 of our book gives psychologically motivated rules-of-thumb for narrowing the choice of formats without resorting to tests.

Organization

The organization of the informational structure refers to the relationship between the different elements of the format. The organization determines how well the particular chosen format is executed. For example, if a format (framework, specifier and labels) chosen is appropriate for the type of information to be communicated but is poorly organized, then miscommuniction is likely. Organization of the elements comprising a format occurs at four distinct levels that we must be aware of - syntactic, semantic, pragmatic and formal.

Initially, we are concerned with how the elements and their constituents treated as a set of marks or lines on a screen organize syntactically. For example, the lines specified must be discriminable, they must group correctly into the appropriate format, and they must not be so complicated the one ca hold them in memory at one time.

Semantic organization concerns the compatibility of symbols and what they are intended to represent. For example, in many process control environments, we are seeing displays that show the process layout using graphic symbols to represent the various components such as valves, boilers, etc. Using symbols that are poor depictions of these components lead to poor organization semantically. Another example of poor semantic organization would be the use of hue to represent a quantitatively changing variable. Hue is a qualitative dimension--physiologically it depends on what kind of neurons are firing rather than how many. Thus, there is an incompatibility between symbol and referent. It would be better, in this case, to use a quantitative aspect of color such as brightness or saturation.

How the informational structure is organized pragmatically determines whether the reader will cone away with information over and above that which is expressly presented. In certain applications, this may be essential. For example, excessive variation in some critical process variable may be unacceptable. If values of this variable are plotted graphically, then expanding the scale should make small variations more quickly noticable and thus lead to better control. Such a display would be organized correctly at the pragmatic level.

Finally, formal organization, taken from Goodman's (1968) theory of symbols, concerns ambiguities that arise when a symbol has more than one interpretation or when a symbol (e.g., label) is missing.

The format and organization taken together define the *configuration* of the display. It is the configuration of the informational structure that drives its interpretation. However, before this structure can be interpreted, it must be made accessible to our information processing system. The accessibility of this structure is made available by the medium which we now discuss.

Medium

Media are the means or devices for conveying informational structures. They are not informational structures in their own right. Examples of media are this page, a TV screen and a telephone. The emphasis in this section will be on a particular medium, the Cathode-Ray-Tube (CRT), since it is becoming the industry "work-horse" and much research has been done on the perceptual effects of its parameters. However, the discussion, where applicable, applies to all other types of display media as well.

For the most part, one is interested in selecting a display device from the multitude of available devices on the market. Given the device, one can then design the configuration appropriate to the task using appropriate software. What is needed, then, is a description of the criteria necessary for selecting the most cost effective display.

Shurtleff (1980) describes two sets of criteria that influence any technical decision regarding the display device. These sets of criteria are:

1. Display design criteria
2. Viewing area criteria

Both criteria will be discussed in turn, and guidelines given by Shurtleff (1980) will be summarized where appropriate.

Display Design Criteria

The basis for both selection and design of CRT's lies in the physical properties of these displays. Shurtleff (1980) has broken these properties into two general categories--classical display properties and modern display properties. The classical display properties are common to all media ranging from printed text on a page to TV screen and include symbol luminance and contrast. The modern display properties emerged with electronic generated displays. They are not as well understood as the classical properties, yet some research exists.

Classical display properties. The design parameters associated with the classical display properties can further be sub-divided into three different types, namely, light parameters, geometric parameters, and display surface parameters.

1) Light parameters. The important parameters for symbol displays are wavelength, lumunance and luminance contrast, with other parameters being direction of contrast and lumunance variations.

The primary perceptual correlate of wavelength is color, which in turn can be specified in terms of its hue (red, blue, etc.), which is primarily related to its wavelength; brightness (light red, dark red, etc.), which is related to the intensity of light; and saturation (a rich full red, a pale red), which is related to the purity of light.

The research summarized by Shurtleff (1980) as well as Kosslyn et. al. (Chapter 3) suggest that for alphanumeric displays, mid-range colors (i.e., green and yellow) should be used. However, colors at the extreme of the spectrum (reds and blues) are suitable for color coding graphics.

Luminance, or intensity of light, has as its perceptual correlate brightness. Simialrly, brightness contrast is the correlate of luminance contrast. Luminance contrast is the relative difference between the luminance of the symbol and its immediate background (see Kosslyn et al, Chapter 3). Luminance variation can either refer to luminance variations within a symbol or variations between two symbols. Its perceptual correlate is non-uniformity in brightness within and between symbols. Finally, direction of contrast refers to the intensity of symbol versus background--bright against dark or dark against light.

The research suggests that the minimum contrast ratio is a function of absolute luminance level and the visual size of the symbols. However, minimum contrast ratios in the ranges of 18:1 allows good symbol identification accuracy over a wide range of symbol sizes and luminances (from 0.1 ft.L to 50 ft.L.). Luminance per se should be kept in the range of 10 ft.L. Direction of contrast is not a critical parameter under most ambient conditions. At extreme ambient conditions, however, a means of changing polarity is desirable.

2) Geometric parameters. These parameters include symbol size, symbol font (shape) and the spacing between symbols, especially in alphanumeric displays. The perceptual correlate of size is the visual angle subtended by the symbol. For alphanumeric symbols, this angle gives rise to judgements of symbol largeness--height and width--and symbol boldness--stroke width. Perceptual correlates of font are not easily determined since shape perception research is equivocal with respect to what is processed--features, templates or prototypes.

Spacing between symbols affects the perceptual distinctiveness of the letters. For example, if spaced to close, individual letters merge together and cannot be discriminated. Spacing refers to both horizontal, within a row, and vertical, between rows. Normally, spacing is measured from center-to-center of adjacent rows or columns.

Design criteria for the symbol height needed for high identification accuracy is dependent on the absolute luminance of the symbol (see also Kosslyn et al (1980), Chapter 3).

For low lumunance levels, 0.01 ft.L. to 0.1 ft.L., a minimum of 20 min. of arc is recommended, for intermediate values of lumunance, 10 ft.L. to 50 ft.L., a minimum of 10 min. of arc is recommended. Also Shurtleff (1980) recommends a stroke-width-to-heigh .tio of 1:5 and a symbol width of 75% of symbol height.

With respect to horizontal spacing, if viewing condition is direct on-line then horizontal spacing can be as close as 8% to 10% of symbol height. For extreme off axes view of 45° or more or low luminance levels, horizontal spacing should be increased to 25% to 50% of symbol height.

3) Display surface parameters. These parameters include the size of the display surface, the display aspect ratio and glare. The size of the display surface is usually stated in terms of the length of the diagonal for retangular displays as a diameter if the display is circular. The perceptual correlate of display size is visual size subtended at the eye of the observer. Aspect ratio refers to the height of the display surface relative to its width. Both these parameters will be discussed more fully in the section on viewing area requirements.

Glare from ceiling lights can completely obscure data from many displays mounted in sloped-front consoles if care is not taken to reduce it. Glare can be eliminated through proper filtering and arrangement of ambient light.

Table 1 summarizes the recommended design parameters associated with the classical factors discussed in this section.

INSERT TABLE 1 HERE

Modern display properties. Shurtleff (1980) divides the parameters that emerged with the advent of the CRT into optical paramters, temporal parameters, electronic paramaters, and symbol generation parameters.

1) Optical parameters. The key optical parameter in CRT displays is focussing of the electron beam. The perceptual correlate of beam defocussing is blurring of the symbols.

The literature summarized by Shurtleff (1980) has shown that a decrease in identification accuracy caused by blurring can be minimized by increasing symbol size and contrast. If this is not possible, defocussing which increases stroke width up to 20% is acceptable. Anything greater than 20% renders the device unacceptable.

2) Temporal parameters. These parameters include fluctuations in the luminance (intensity) of the electron beam, the perceptual correlate of which is flicker. The length of time for a light output to decay to some percentage of initial brightness is called the persistence, and this parameter varies from phosphor to phosphor. To keep the light output at a constant level to the perceiver, it must be updated or 'refreshed' at a rate above the flicker-fusion frequency of about 40 Hz. Update rates less than this critical value will cause flicker--light appearing to fluctulate in intensity at approximately 100 msec intervals. If display refresh is less than 40 Hz, then it is possible to compensate for flicker by using a longer persistance phosphor. Flicker should be minimized at all times since it can overly stress the visual mechanisms resulting in visual fatigue.

3) Electronic parameters. These parameters are characteristics of electronics hardware that change the positioning and stability of the electron beam. For example, noise and instability in the amplifiers and other driving circutry that involve a gradual change in the position of the electron beam lead to the perceptual phenomenon known as drift. If the change in position is small, abrupt, relatively quick and repititious then it is perceived as jitter.

Shurtleff (1980) states that while these parameters produce annoying perceptual sensation, they are not critical to symbol identification. Nonetheless, their effects on the task at hand should be assessed prior to device selection.

4) Symbol generation parameters. While there are a multitude of CRT terminals on the market, making selection of a terminal appear almost impossible, they all basically fall into one of three major technologies for producing graphics--raster scan, stroke or calligraphic, and direct view storage tube (DVST). Each technology has its pros and cons as well as intratechnology considerations that we should be aware of before a final selection is made.

The DVST behaves like a CRT with a long persistance phosphor. Persistence is defined by the time it takes light output emitted by the phosphor to decay to some percentage of its original brightness. Thus when the electron beam hits the phosphor backplate, the affected phosphor retains a glow for hours after the 'hit'. In this technology, line segments are drawn by directly moving the electron beam from point to point. Once the pattern is traced, the charged phosphor results in a relatively permanent trace on the screen because of its long persistence.

The advantages of DVST technology are its high drawing precision and lack of screen refresh. Since the system draws lines directly from point to point, precision is limited on by the focusing capability of the beam. Thus line qulaity can be as good as ink drawings. Since no refresh is necessary, flicker is eliminated from these displays.

This lack of refresh capability also imposes a constraint on the system. It cannot be used for real time applications. The whole screen must be erased and the picture must be completely redrawn to update it. Additionally, another disadvantge with the technology is that it is essentially a one color medium.

Color capability and gray scale imaging are not presently available with this technology. Prices range anywhere from $3,000 to $20,000.

Stroke writing or calligraphic sytems create vectors in much the same way as DVST systems, direct movement of the electron beam from point A to point B. However, the phosphor has a quick persistence requiring refresh. This allows real time updating of the picture. However, the refresh rate, usually 40 Hz, limits the amount of data that can be displayed on the screen at any one time. Since the beam moves at a relatively constant rate, the longer the line, the longer it takes to draw. Thus flicker is a real problem if the line is too long.

In alphanumeric displays, the parameters related to stroke element design or selection are techniques used to generate the symbols; and the matrix size used to construct the symbol. Figure 1 shows examples of fixed positions and random position matrix of strokes for both 5 x 7 and 9 x 9 matrices. Shurtleff (1980) has summarized research showing that random position stroke writers lead to higher identification accuracy than fixed position stroke writers. With regard to matrix size, at a minimum, the size should be 5 x 7 with the 9 x 9 preferred if degradation in viewing conditions is anticipated.

Although color is not readily available with stroke writing systems, it is becoming more common for manufacturers to have stroke writers with limited color capability. Differing color and speed requirements serve to give this technology the widest price range--anywhere for $10,000 to $100,000.

INSERT FIGURE 1 HERE

The raster scan CRT is the most familiar display type. In these displays, an electron beam sweeps horizontally across the screen from left to right, drawing the picture as a series of scan lines. After the end of each line, the beam is turned off and repositioned at the beginning of the next line. This process continues until the beam finishes the last line, at which point it is repositioned at the top and the process repeats itself. Since the picture has to be refreshed, flicker is also a consideration with these displays.

Logically, the screen is divided into thousands of small regions called pixels. Raster displays draw lines as a series of pixels, which make diagonal lines appear as a staircase and, consequently, limits drawing precision.

Spatial resolution of a raster scan system is given as the product of the number of horizontal pixels and the number of vertical pixels or horizontal scan lines. Color and gray scale resolution, common in raster scan graphics, is determined by the number of color/intensity choice at each pixel. Thus, for example, a raster scan system that has 256 x 256 by 4 bit resolution can present 256 squared resoluble pixels, each of which may display $2^4 = 16$ possible colors/intensities.

The information on a raster screen is refreshed from a 'bit map' where each pixel corresponds to a group of bits in memory. The simplest bit map display assigns a single bit to each pixel. This would correspond to a color/intensity selection of two--black and white, for example.

For alphanumeric display, one is often concerned with the vertical resolution or number of scan lines per symbol height. The design guideline derived from the literature is that a minimum of 10 or 12 scan lines per symbol height are needed for maximum speed and accuracy of identification.

Raster scan terminals are usually less expensive than stroke writing systems, in the range of $3,000 to $6,000. Table 2 shows the comparison of the different graphic display technologies.

INSERT TABLE 2 HERE

1

Finally, one last symbol generation technique for alphanumeric displays must be discussed--dot matrix display. This method of forming symbols uses a matrix of dots, ranging from as few as 35 (5 x 7), which is the industry standard, to as many as 140 (10 x 14). Dot matrix size is the key parameter affecting accessibility of symbols generated from this display as Figure 2 indicates.

INSERT FIGURE 2 HERE

With respect to this parameter, design guidance suggests that the standard 5 x 7 matrix may be used when display quality is good. If degraded viewing conditions are anticipated, then matrices of 7 x 9 or larger are recommended.

Viewing Area Criteria

Shurtleff (1980) has developed criteria for viewing area requirements that represent a necessary adjunct to the design criteria specified in the previous section. The criteria guarantees wide angle viewing by insisting that a reader be able to identify symbols when working at any position on the display shelf, and working at a close-in position, or in a relaxed position. Figure 3 shows a viewing area defined in this way. The required viewing area is shown as a rectangular area extending from the outside edges of the console shelf to 14 inches beyond the console itself. This area constrains both the symbol height and size of the CRT, since symbols on the right side of the screen must be large enough to be seen at the extreme left side of the required viewing area and vice versa.

INSERT FIGURE 3 HERE

The procedure, detailed in Shurtleff (1980) begins by determining the maximum on-line viewing distance and minimum display size needed to encompass the viewing area requirements of a given problem. This is followed by the determination of the appropriate symbol size necessary to acheive 95% identification accuracy at the required viewing distance. Following this, one has to determine whether the symbol size mandated implies a larger screen size due to the capacity requirements (i.e., the number of horizontal and vertical symbols) dictated by the problem. Finally, one has to determine the resolution requirements of the system given symbol size and capacity requirements.

Summary

Both components of the display design problem--medium and informational structure--plus their associatied design issues have been presented. Together, these two components provide the framwework for specifying both the hardware and software requirements for any graphics display system.

We now turn our attention to a procedure for applying this framework in an orderly fashion to any display design problem we might encounter.

PROCESS

Having developed a framework for analyzing the display design problem, we must now apply the framework to a particular design problem. Since the solution relies of technological factors, the resulting product is said to have been engineered or designed. It is the purpose of this section to describe the process of design and show how it can be used for the particular problem of display design. To begin, though, I will give a brief overview of the context in which this process has evolved.

Background

Since design involves problem solving, early characterizations of the design process described it in terms of the procedures for solving problems. For example, Buhl (1960) describes the process of design in terms of the following steps:

Recognition of the problem and the decision to do something about it.

Definition of the problem specifically, in familiar terms and symbols; delineation of the problem into subproblems and goals, placement of the necessary limitations and restrictions.

Preparation by compiling all past experience in the form of data, ideas, opinions, assumptions, and the like.

Analysis of all preparatory material in view of the defined problems and evaluation of all information which may bear on the solution.

Synthesis of a solution from the analyzed information. Assembling the various items analyzed to produce possible solutions.

Evaluation of possible solution and selection of the best solution. Verification and checking of the many aspects of the solution and the integration of all sub-problem solutions into a unified whole.

Presentation of necessary information to others in order to execute the solution. Actuation of the solution to satisfy the recognized need.

In 1962, Morris Asimow used this problem solving process to develop a philosophy of engineering design. Asimow (1962) defines engineering design as a systematic activity directed toward the goal of fulfilling human needs using the technological factors of our culture. Asimow's (1962) philosophy encompasses three major parts; a set of consistent principles and their logical derivatives, an operational discipline for executing the design, and a feedback mechanism for guaging the effectiveness, detecting shortcomings, and illuminating the directions of improvement of the design.

Asimow (1962) embeds this philosophy in the production-consumption cycle since the products designed must enter this cycle and must be compatable with its processes. These processes shown in Figure 4 are production, distribution, consumption, and recovery or retirement. Each of these processes places its own demands on the design, some of which may be contradictory implying trade-offs. For example, the designer is basically concerned with the consumer, for without consumer's satisfaction with the product it will fail. But he must also concern himself with the needs of the producer who is also his employer. While the demands from the consumer will be couched in terms such as reliability, ease of maintenance, aesthetics and the like, the producer demands ease of production, standardization of parts, availability of resources and with reduction of rejections. The distributor in turn demands ease of transport, suitability of storage, long shelf life and so on. Finally, in retirement we must be concerned with the reusability of parts and ecomonic disposal of the remainder. The designer must factor each of these separate and sometimes divergent viewpoints into the design of products.

INSERT FIGURE 4 HERE

In order to prepare a product for entrance into the production-consumption cycles, Asimow (1962) sets forth a chronological sequence of primary design activities shown in Figure 5. These activities are the feasibility study, preliminary design, and detailed design. The purpose of the feasibility study is acheive a set of useful solutions to a recognized need. The preliminary design selects the best design concept from this set. Finally, the detailed design activity furnishes the engineering description of a tested and producible product.

INSERT FIGURE 5 HERE

Each of these primary design activities can, in turn, be broken down in its constituent steps. In the remainder of this section, the process of design defined by these three major design activities is broken down into the constituent steps of each and applied to the display design problem.

Feasibility Study

The feasibility study is the first step in the design process, the purpose of which is to identify potential solutions to the design problem. It does so by seeking answers to the question "What is the problem?" and "How can it/be solved?" Note that the second question should be asked only after a careful analysis of the first. That is, preconceived solutions based on quick preliminary analyses should be avoided since they tend to limit the creativity of the process. To many times a designer has a "pet" concept for solving the problem and follows it blindly without considering any alternatives which may prove more suitable. This biased approach to design can be eliminated if we follow the structured steps that make up the feasibility study, shown in Figure 6. The first two steps, needs analysis and derivation of display system requirements define the problem in a very precise technical way, allowing a technological solution. The final two steps, formulations of the design concept and synthesis of solutions begins the process of technological realization. Note that the design concept is formulated only after the problem is understood to such an extent that display inputs, outputs and constraints can be specified.

INSERT FIGURE 6 HERE

Needs Analysis

The display design process begins with a recognition of needs, real or latent, that must be satisfied. The purpose of the needs analysis is to determine these needs to such an extent that a basis for subsequent design decisions is formed. In order to accomplish this purpose, we must clearly define the design problem and orient ourselves in the design space. That is we, as designers, must become educated in the different ways of approaching the problem.

Statement of purpose. The first step of the needs analysis is to state the purpose of the design. The purpose is derived from some primitive statement of needs, primitive in the sense that the needs represent educated opinions based mainly on observation but, as yet, unsupported by systematic data. Depending on the particular environment in which the needs originate, they can be framed in terms of a system's mission (e.g., to monitor and display information regarding...), goal (e.g., to develop a particular graphics presentation to emphasize the current state of company A's financial position) or simply a need itself (e.g., a display system is needed which...). It is the responsibility of the designer to evaluate the feasibility of the need and to evolve a precise definition.

Pitts (1973), in his process of design, has set forth some points to be considered in evaluating needs and deriving a statement of purpose. The first point to be considered is the reliability of the source from which the need originates. Sources can be internal or external. Internal needs originate within the designer's organization in one of two ways, either as a subcomponent of a larger design or a new product. Pitts (1973) points out that all to often subcomponent requirements of larger designs are not recognized as designs in their own right but yet should be. External needs are provided by a customer who is able to specify the need in a general way but may not have the specialist knowledge to define it exactly.

Thus, as a second point when a doubtful need arises, it must be initially justified on economic grounds. Sometimes the customer is unaware that a suitable display exists, meaning that the new display must have some competitive edge if it is to be successful and, hence, warrant further design consideration. Similarly, in a limited marketplace, the cost of producing a product must fall below the specified maximum price if the need is to be justified. The cost of the product must be viewed within the entire production-consumption cycle. A breakdown of product cost within this framework is shown in Figure 7 and should serve as a useful guide.

INSERT FIGURE 7 HERE

Finally, the statement of purpose should be so framed as to avoid unnecessarily influencing the possible range of solutions. For example, a customer may have stated the need for a color graphics medium when the real need is for a graphics medium. By stating color, the customer has implicitly made a design decision. The designer should translate this need into a statement of purpose of, say, the following: form of visual presentation.... Only after the input/output requirements have been specified should decisions regarding the use of color be made. At this stage in the design process, nothing that in any way limits creative thinking should be allowed.

Orientation. After the statement of purpose has been explicitly set forth, we should begin the process of orienting ourselves to solving the problem. The designer does not usually work in a vacuum. There are many different ways for obtaining useful information relevant to the design.

1) Talking with the client. A useful starting point in the process of gathering and assimilating information is the customer himself. Most clients have a fairly good idea of what the relevent issues are, the likes and dislikes of a particular group of users, as well as knowledge of industry specific journals that can help familiarize the designer with the competition's products and the like. Also, talking with the client can identify key internal resources which may prove useful to the designer.

2) Background search. Additionally, there may have been previous analyses of this problem or similar problems which can yield very useful information about critical parameters or key system requirements. In fact, the gathering and searching for this background material constitutes the initial phase of Kreifeldt and Hill's (1974) methodology for tool design. They list three primary modes of search--patent search, literature search, and morphological ordering. Morphological ordering involves gathering exemplars of the potential design and arranging them for visual analysis. With displays, exemplars will be pictures and design specification. The arrangement should highlight commonalities within a group and differences between groups. Such a technique allows relevant design parameters to emerge as well as suggesting certain functions that must be performed and the existing means of performing them.

3) Survey. Thus far, the orientation effort has, for the most part, provided information concerning the display itself. However, these displays must

be used by people in performing certain tasks outlined in the previous section. Therefore, a complete view of the design problem must involve information on user, display, and task and how they interact. Additionally, the background s[illegible]. may have left 'holes' that must be filled by information. Information about the design from a user oriented viewpoint can be obtained by surveying or polling the user population. Such techniques are intended to elicit users conception of the task, on how it should be performed, on how the display actually assists in performance of the task, as well as likes and dislikes concerning existing competitive displays.

4) Contextual description. Many times a display is embedded in a much larger system of activities. Describing this set of activities or context can help the designer by providing a structure to the design problem. This description can take on many forms, such as an operational flow diagram, mission profiles and the like.

Derivation of System Requirements

The needs analysis with its accompanying orientational activity should provide the background necessary to specify the requirements the display system must satisfy. The purpose of this step is to subjectively bound the design requirements in order to evolve the most effective solution that satisfies these requirements. These requirements provide input to the selection and design of both the informational structure and the medium. Following Ostrofsky (1977), this is accomplished by considering all phases of the production-consumption cycle in terms of inputs--both endogenous or system induced and exogenous or environmentally induced--and outputs--desired and undesired. There are two types of system requirements that must be considered--physical and operational-behavioral.

Physical. These are the technological requirements that affect the product design and selection of the medium. These requirements are directly quantifiable, usually as a result of calculation. They include such things as reliability, cost, as well as the performance requirements specified by criteria in Tables 1-3.

Operational-behavioral. These are the requirements that relate to the operational environment that deals with human task capability. These requirements are usually not directly quantifiable but may be described in terms of characteristics that bear some monotonic relationship to them. For example, a display must be designed to minimize the cognitive load on the operator. Operationally cognitive load could be indicted by response time to a secondary task. Thus, using this criteria, it is possible to assess different display configurations under experimental conditions.

The environmental requirements are usually framed in terms of the conditions under which the display will be used (Grether & Baker, 1972). These conditions are in part derived from the viewing requirements and include:

1) Viewing distance. This requirement influences the size of the displayed symbols as noted in the previous section.

2) Illumination. The size of the displayed symbols should be suited to the lowest expected illumination. Also when using general illumination for display, glare considerations are warrented.

3) Angle of view. This requirement also affects symbol size.

4) Presence of other displays. An operator usually divides his or her attention among several displays. The displays must be grouped to allow maximum scanning efficiency (another operational-behavioral requirement).

5) Compatibility with related controls. The type of control action and placement of control should be easy to locate and use.

A useful aid for deriving system requirements is an input-output matrix for each phase of the production-consumption cycle with narrative descriptions for each respective cell in the matrix. A group of such matrices, used by Ostrofsky (1978) in designing a power unit support stand for the F-16 Aircraft, is shown in Figures 8-11. Note that the descriptions are general enough to be applied to most equipment designs. However, as Ostrofsky (1978) notes, the exercise in filling out these input-output matrices induces considerable awareness amongst designers of the major design problems that must be investigated. He also notes that operational-behavioral requirements arise naturally from these considerations.

INSERT FIGURES 8-11 HERE

Alternatively, requirements may be set by the customer and are found in such documents as the request for proposal (RFP), the statement of work (SOW), and any preliminary test reports. These requirements, together with those arising from the input-output activity will be used to evaluate the subsequent design.

Formulation of Design Concept

The design concept is defined by Ostrofsky (1977) as a basic approach toward solving the requirements problem. It is simply a delineation of the display formats and organizations and medium properties--classical and modern--necessary to allow the designer to identify alternative display configurations and mediums for realization.

Beginning the process of concept formation requires knowledge of the system inputs and outputs. These in turn require certain functions for their implementation. They may also suggest certain functional capabilities or impose performance requirements and constraints. One usually works backwards from the performance goal performing a functional analysis.

Functional analysis. This method of establishing a design concept is an explicitly defined requirement for work on military systems, equipment and facilities (see MIL-H-46855B). It should likewise be part of every designers proposal for commercially defined systems [illegible] products as well.

Following Meister's (1971) design methodology, one starts the analysis by listing sequentially the individual major activities or functions necessary to implement the goal. Again, as stated before, how detailed this breakdown should be will depend on the nature of the design itself. An extremely useful way of describing the resulting functions is in the form of a functional flow block diagram (FFD). The advantages of such a diagram are two-fold. First, determining the logical inputs and outputs for each function effectively defines the sequence in which these functions should be arranged. Secondly, if we can determine the time required to perform each function, we can plot the FFD on a time continuum and compare this with any relevent performance requirements. An example of a functional flow diagram for identifying and transporting spent-fuel canisters emplaced in salt to some surface recieving station (Harris, 1982) is shown in Figure 12. Note that between identification and transportation, additional functions are required, and they are inserted into the diagram. Note also that feedback loops are inserted where necessary.

INSERT FIGURE 12 HERE

Once the FFD has been described, determine the system and environmental constraints that impact each function. These are usually listed in a tabular narrative format as shown in Figure 13 for the canister removal function illustration.

INSERT FIGURE 13 HERE

In complex systems where decisions and actions are part of the functional flow, they should be represented by decision/action diagrams which add to the description of system requirements. Since decisions are a vital contribution of the human to th cess of the mission or goal, they are important for human factors analysis. A decision action diagram for the Survey Function of the above illustration is shown in Figure 14.

INSERT FIGURE 14 HERE

In general, the procedure for building up on PFD is to analyze each system function in terms of its required inputs and outputs. Each of these are in turn analyzed in a similar way to elaborate the functional description. At this point in the overall design, however, these inputs and outputs do not imply the actual means by which they are produced. This comes at the next stage.

Synthesis of Solutions

Synthesis is the process of fitting together the seperate functional concepts into an integrated whole--a particular informational structure and medium. As Asimow (1962) notes, it is the synthesis step that clearly characterizes the project as a design undertaking. This step is the creative part of the design process. That is, ideas must be generated, ideas that represent alternatives to each equipment or system function. A display system that arises as a consequence of a specific set of alternatives is called a <u>candidate system</u> (Ostrofsky, 1977).

Alternatives, and ultimately candidate systems, come from stimulating ones imagination. It is the primary responsibility of the designer to be innovative in the choice of alternatives, and while the creativity for arriving at innovative solutions must come from the designer himself, techniques exist for enhancing the creative attitude.

Creative synthesis of candidate systems. Following the design philosophy of Alger & Hays (1964), two general classes of methods can be used to arrive at a number of potential candidate systems. These are individual creative effort, and group creative effort

In the first method, a particular path of logic is developed and used to expand the alternatives. The second method relies on the interaction between members of the group to stimulate ideas.

1) Individual creative effort. One technique for enhancing the creative attitude of the individual is referred to by Alger & Hays (1964) as morphological analysis. This technique is a very systematic and orderly means of generating innovative candidate systems by forcing an association of a few basic product or system functions. First, the major functions of a product or system must be established (this has been done in the functional allocation stage of the design process). Second, the means of performing these functions must be conceived by the individual through an ideation session. Next a matrix is set up with each function being a dimension of the matrix. Values on each dimension correspond to different methods of performing the function. For example, they could be the different graph types defined by Kosslyn in Chapter 7 of our book. Thus each cell of the matrix represents a particular candidate system. Another level of morphological analysis can be carried out for specific candidate systems to detail the design concept more completely.

As an illustration of this technique, consider the design of a clothes dryer taken from Alger & Hays (1964). The first step is to delineate two or three major functions required to perform clothes drying. These are 1) Source of heat, 2) Environment around clothes, and 3) Drying mechanism. Each of these functions constitutes a dimension of a three dimensional matrix, the values on each dimension being the different ways of implementing each function. This is shown in Figure 15. Each combination of values along the dimensions defines a candidate system. Next, a particular candidate system can be analyzed in more detail as shown in Figure 16 for the candidate system described by the gas environment electric tumbler dryer.

INSERT FIGURES 15 & 16 HERE

In summary, morphological analysis allows the delineation of a great many candidate systems from a few basic functions of a product or system. Note that the innovativeness of the design rests on the designer since it is he who must derive the different means of implementing each function.

2) Group creative effort. This form of ideation has come to be known as brainstorming. The key to brainstorming is a relaxed atmosphere in which the participants are encouraged to suggest as many solutions as possible no matter how far fetched they might be. Judgement of the worth of any idea is withheld.

Alger & Hays (1964) set forth some rules that help ensure a successful session. The rules are as follows:

- State the problem as simply as possible.
- Rule out judgement. In a formal session, the group leader should stop any member who offers an evaluation. Evaluation must wait.
- Have members mention all ideas. The wilder the idea, the better. It is usually easier to tame down an idea than to build it up. Someone else may suggest a change which makes a previously impractical idea successful.
- Encourage members to give as many ideas as possible.
- Encourage members to combine and improve ideas suggested earlier.
- To vary the pace, encourage humor to relax the participants.

- Limit the session time to a period stipulated at the beginning. Half an hour to one hour maximum is usually appropriate.
- Assign a recorder to take down the gist of each idea and the originator's name on a blackboard or large chart pad in front of the participants. Keep the list of ideas in front of the participants during the session.
- Write up the ideas carefully after the session and send them to each participant. Encourage them to jot down others during the following two or three days.
- Summarize the total output of ideas by grouping them under logical headings. Normally major headings are the basic design-concepts.

Brainstorming appears most helpful in stimulating ideas for new production or new product functions. Thus, a group brainstrom should probably be reserved for more important design problems and at early stages of the design. However, individual brainstorming is always recommended.

Screening of candidate systems. The outcome of either type of ideation session should be a number of candidate systems. Obviously, it is not possible to evaluate the feasibility if the number of systems is large. Thus some sort of screening process is necessary if we are to reduce the list to a few desirable candidates. This screening process probably should occur in stages.

In the first stage, candidate systems are eliminated on the basis of common sense or noncompatibility between functional values. For example, certain cells arising from a morphological analysis may have combinations of values that are mutually exclusive. Also, certain candidates arising from a brainstorming session may simply be too far fetched to be very fruitful design alternatives.

The candidate systems passing this initial screening are then evaluated on their technical feasibility. At this stage, the designer is merely interested in orders of magnitude to determine if certain candidates are technically out of the question.

This screening phase is followed by an evaluation based on economic considerations. These economic considerations include economic worthwhileness and financial feasibility (Asimow, 1962). Economic wothwhileness concerns the 'utility' of the product to the user. That is, will the product as developed be of sufficient value to the consumer that he will be willing to pay for the effort to produce it. This is exceedi y difficult to answer and requires the designer to place himself at the three points of the production-consumption cycle--producer, distributor, and consumer--and evaluate each candidate system from each of these perspectives.

Finally, the designer must evaluate each candidate system in terms of the financial resources necessary for its realization. If a particular candidate system is too costly, it should be abandoned.

Like the saying "many are called but a few are chosen", the screening process starts with a large number of potential design solutions but, if applied correctly, allows only the most useful of solutions pass to the next stage in the design process. Usually, these solutions are so close on technical and economic merit that more formal approaches are needed to select the 'optimal' candidate.

Preliminary Design

The set of candidate system developed in the Feasibility Study must now be analyzed in terms of their design parameters and the best alternative chosen. This is the purpose of the preliminary design phase. The steps comprising the design activity are shown in Figure 17.

INSERT FIGURE 17 HERE

Selecting the best alternative requires that we have some criteria for deriving the best. Thus the first step in this activity is the identification of formal design criteria against which the set of candidate systems can be

3x

evaluated. Next we must analyze each criterion in terms of its relevent design parameters. The description resulting from such an analysis should be in the form of equations relating the parameters to criterion. These equations must then be combined into a criteria function which will provide an overall figure-of-merit for each candidate system. An analysis of the design space defined by the design parameters should be undertaken with respect to the criteria function. This analysis provides the designer with insight into the design space by allowing him or her to gauge the impact of variations in design parameters on system performance. Thus for the particular candidate system chosen, the tolerances of its parameters will be known well ahead of any realization effort eliminating any surprises during testing. Finally the best candidate system must be chosen. This is the optimazation step.

Criteria Specification

In order to evaluate the performance of the set of remaining candidate display designs, criteria must be explicitly defined. In most cases, these criteria can be identified from the requirements analysis (i.e., the input-output matrices) and any accompanying documentation. In other cases (cf. Ostrofsky, Donaghey, Marquina & Kiessling, 1980) a questionnaire can be developed and given to the client to elicit criteria.

The criteria identified are sufficiently abstract that I will give them a special name--design constructs. These constructs concern the domain of both physical and operational-behavioral variables and include criterion such as producability, reliability, durability, ease-of-use, aesthetically pleasing, cost, safety and the like. In a later design step these constructs will be given a more operational definition in terms of measurable design parameters.

Weighting the criteria. Applying the criterion to the selection of a candidate system requires the designer to have some idea of the importance or weight each criterion should contribute to the overall function. For example, in some systems or products, cost may be the overriding factor, in others it may be reliability. In still others it may be a group of criterion.

The determination of the various weights each criterion should have is based on a subjective assessment by both client and designer. However, formal procedures exist for helping to quantify this judgement.

For example, a procedure summarized by Meister (1971) is shown in Table 4. Weights are assigned by comparing each criterion with every other and assigning a 1 to whichever of the two is judged more important, a 0 to the criterion judged less important. This process is continued until the entire table is filled. The 1's in each row, corresponding to each criterion, are then summed together and divided by the total number of 1's assigned to give a weighting coefficient for that criteria.

INSERT TABLE 4 HERE

Alternatively, the questionnaire developed by Ostrofsky et al (1980) allowed each respondent to rate the relative importance of each criterion on a scale of 0 to 10, ten being the highest rating. The mean ratings for each criterion were then scaled by the overall mean to give an indication of relative importance for each criterion.

Criterion definition. Now that we have identified the design constructs and weighted their importance, we must analyze these constructs in terms of the variables of the design equipment and any environmental variables that impact the operational ability of the equipment. These variables serve to identify each construct explicitly for the optimization step. We obtain the insight into which variables are relevant to the design in question from the steps of the feasibility study (e.g., orientation).

These design variables generally fall into three descriptive categories: directly measureable, indirectly measurable, and observable.

1) Directly measureable variables. These are display and environmental variables that are directly quantifiable. It is these variables, called *design parameters*, whose specification ultimately dictates the design in question.

2) Indirectly measureable variables. These are variables, called *submodels*, whose values arise as a result of relationships existing between the design parameters. However, once arrived at, values of these variables can be given a precise meaning. For example, the probability of detecting a signal in a display will depend among other things on the intensity of the signal. Thus if the intensity is a design parameter, probability of detection is an indirectly measurable variable.

3) Observable variables. These variables are not directly quantifiable but yet have an impact on the design. Their effects must be tested empirically. Examples of these variables would be the arrangement of controls, control accessibility as well as display configurations. A method of quantifying the accessibility of controls had been recently prepared by Banks and Boone (1981). This index makes use of the operators reach envelope, the frequency of use of each control and the position of each control relative to the operator. This is a step in the right direction and more research is needed to establish the reliability of such indices.

Following Ostrofsky (1978), the way in which the criteria is defined proceeds as follows. For each construct identified list all the variables that relate to it. Next designate each variable as directly measureable, indirectly measureable or observable. Finally, in a hierarchical arrangement, from construct to submodel to parameter, indicate which variables relate to each construct. An example of this approach for the power unit support stand designed by Ostrofsky (1978) is shown in Tables 5 and 6.

INSERT TABLES 5 AND 6 HERE

Modelling the Criteria

The modelling process begins the exercise of precisely defining the relationships between construct, submodel, and design parameters. These relationships take the form of mathematical equations and are guided by analytical creativity.

The process begins by listing all assumptions, submodels (z_j), parameters (y_k), and givens (c). For each construct X_i, there will be some direct functional relationships between the contruct and design parameter represented as

$$X_i = h_k(y_k) \tag{1}$$

In the absence of any direct relationship between design construct and parameters, an indirect functional relationship can be derived through the use of submodels as follows. There is a direct relationship between construct and submodel

$$X_i = f_i(z_j) \tag{2}$$

Similarly, there is a direct relationship between submodel and design parameter

$$z_j = g_j(y_k) \tag{3}$$

Therefore,

$$X_i = f_i\left[g_j(y_k)\right] \tag{4}$$

The resulting model relating design construct to design parameters can be expressed as

$$X_i = f_i\left[g_j(y_k)\right] + h_k(y_k) \tag{5}$$

This formal design step, advocated by Ostrofsky (1977), requires extensive technical competence on the part of the designer and in most cases requires a collaborative effort from a number of technical disciplines. In complex projects this collaboration can reduce communication between the various members of the design team.

Synthesis of Criterion Function

Formal optimization requires the designer to synthesize a criterion function which is used to evaluate each candidate system. That is, the weight, a_i, and criterion, X_i, must be combined into a single expression representing a figure-of-merit for each candidate system. This expression is then used to assess the performance of a given candidate system resulting from a particular combination of the design parameters y_k.

Parameter range specification. In order to obtain the minimum and maximum values of each criterion, we must specify the range of each of the design parameters and through the functional equations derive the criterion range. It is very important that meaningful values for the ranges be supplied since these values are used to help define the overall criterion function. Thus if the ranges are not accurately defined we may delete desirable alternatives or add an unduly large number of undesirable alternatives.

Normalization of critera. Since each criteria may be defined on a different scale some method of acheiving comparable units is necessary. Without this standardization, the resulting composite function would be meaningless. Following Ostrofsky (1978) his standardization is accomplished by identifying criterion performance as a fraction of the allowable range for that criterion.

$$x_i = \frac{X_i - X_i\text{min}}{X_i\text{max} - X_i\text{min}} \qquad (6)$$

where $X_i\text{max} - X_i\text{min}$ is the criterion range derived previously. This expression also allows us to gauge the sensitivity of a unit change in, say, X_1 with a unit change in the other criteria.

Synthesis. When the fraction of the criterion range that a given candidate system will yield as its performance, as indicated by equation (6), is combined with its weight a_i, the expression $a_i x_i$ represents the relative value of the i^{th} criterion. The relative values for each criterion must then be combined to yield an overall predictor of performance--the criterion function (cf). This combination can occur under 1) assumption of criterion independence, or 2) that interactions among the criterion exist. I will consider each in turn.

1) Criterion independence. When the criteria are independent, the criterion function can be expressed as

$$cf = \sum_{i=1}^{m} a_i x_i \tag{7}$$

or in terms of the design parameers y_k as

$$cf = \sum_{i=1}^{m} a_i \frac{\left(X_i - X_i \min \right)}{\left(X_i \max - X_i \min \right)}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{m} a_i \frac{\left[f_i(z_j) - X_i \min \right]}{\left[X_i \max - X_i \min \right]}$$

$$cf = \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{p} a_i \frac{\left\{ \left[f_i(g_j(y_k) + h_k y_k \right] - X_i \min \right\}}{\left\{ X_i \max - X_i \min \right\}} \tag{8}$$

Thus, for a given set of y_k's, equation (8) will return a single number which is then used as the figure-of-merit for that design alternative.

2) Criteria interaction. When the criteria interact, probability considerations are necessary to synthesize the criterion function. That is, each x_i is embedded in a probability space with distribution $F(x_i)$. Following Choudhudy (1979) the criterion function in this case becomes

$$cf = \sum_{i=1}^{m} a_i F(x_i) - \sum_{i<j} a_{ij} F(x_i x_j) + \dots + \sum_{i<j<k} a_{ijk} F(x_i x_2 \cdot\cdot x_n) \qquad (9)$$

where $F(x_1 x_2 \cdot\cdot x_i)$ represents the joint distribution of the constructs and $F(x_i)$ represents the marginal distribution. Note that since the criterion function is a probability measure, it is defined on the interval zero to one. An approach for obtaining the marginal and joint distributions of these constructs is also presented by Choudhury (1979).

Analsis of Design Space

Now that we have derived the criterion function we are in a position to analyze the design space. This space is defined by the number of free design parameters[1] plus the criterion function, the limits of which are set by the ranges of the parameters and the limits of the criterion function, zero and one.

Sensitivity analysis. Before we implement any specific display system, we would like to know how sensitive the criterion function will be to changes in the design parameters. This will allow us to specify the tolerances necessary to ensure a relatively stable criterion function during actual implementation. As Asimow (1962) notes, the results of the sensitivity analysis provide the designer with a deeper insight into the inner wakings of the product or system; an indication of the critical design parameters; an indication as to whether some of the constraints should be relaxed or tightened; and a better feel for the quantitative performance of the product.

---

[1] In some designs certain of the design parameters may be held constant for all candidate systems. The values are usually specified in the accompanying documentation.

Sensitivity analysis involves the determination of the rate of change of the criterion function for a given rate of change of each design parameter throughout its range. This procedure is invariably performed numerically through software.

Compatibility analysis. In large scale systems with many displays, individual displays must be used as a group to perform the overall task. If this overall performance is to be successful, these individual displays must be compatible with one another.

Determination of compatibility may involve the assessment of tolerances (e.g., geometrical, physical). However, more difficult problems of compatibility arise when the human is considered as a subsystem whose characteristics must be matched to that of the equipment. Human Factors considerations (cf. Van Colt & Kinkade, 1972) are necessary to ensure the likelihood of mismatch is minimal. If a mismatch is found, note that the parameters exhibiting the least critical effect from the sensitivity analysis can receive major changes towards accommodating the system to enhance compatibility.

Optimization

In order to proceed with the final phase of the design, the optimal candidate system must be identified. This requires the criterion function defined by either equation (8) or equation (9) to be evaluated for each candidate system and that candidate system yielding the highest value is selected.

Following this determination, Ostrofsky (1978) recommends freeing all design parameters and re-evaluating the criterion function by searching the design space. The system that results can then be thought of as yielding an upper bound in performance ~~that~~ may never be acheived in practice. Nonetheless, it represents a target that future iterations in the design can approach.

Detailed Design

It is in the detailed design stage that the product as a piece of hardware is realized. This realization is acheived by constructing a prototype from a set of design instructions and represents phase V of K ldt & Hill's (1974) design methodology. However, like the other design phases, this phase can be broken down into its constituent steps as shown in Figure 18.

INSERT FIGURE 18 HERE

Initially, resources are allocated to the project in preparation for design. This is followed by the design and assembly of the prototype. This prototype is then subjected to test and evaluation followed by a redesign effort if necessary.

Allocation of Resources

Figure 8 from the needs analysis section showed the breakdown of costs associated with the design and production of a product. The estimates made during that stage were quite crude--a first pass estimate necessary to flag economically excessive projects in their tracks before large commitments were made. At this stage of the process, the design concept has been accepted and large scale commitments are necessary. Thus the precision of estimates at this point must be relatively high to ensure a sound basis for subsequent decision making.

These estimates are framed in terms of time-to-completion, manpower necessary, and support equipment which taken together form a pre-production budget request.

Time-to-completion. The responsibility falls on the designer for deciding the length of time necessary for a project. The basis for this decision should include:

1) Final cost of the product
2) Demand for the product
3) Number of people involved
4) Wage rate of the pec
5) Material procurment delays
6) Overhead

Additionally, this decision should be based on the designer's intuition concerning:

1) Task complexity
2) Experience and caliber of the people involved
3) New knowledge or techniques that must be learned
4) Incomplete information

Scheduling aids such as PERT--Performance Evaluation Review Technique--(cf. Hill, 1970) can be used to systematically develop a work plan. PERT does this by requiring the designer to breakout the tasks required, list them in a chronological sequence, estimate times required to complete each, and then add up all the times necessary to complete the project.

Manpower. For designs of sufficient complexity, a group effort will be necessary. This requires individuals with different technical skills and specialties to be brought together for the common purpose of producing a product. Though many organizational alternatives exist (cf. Meister, 1971), they are all based on the following assumptions:

1) Since the major role of the specialist is to assist product development, whatever facilitates this assistance improves design efficiency.
2) Any direct contact between specialist and engineer improves the effectiveness of this assistance and consequently improves design efficiency.

3) Any organization that places the specialist and engineer in most direct contact will be most effective.

Since the primary detriment to any design project is lack of communication between the members of the team, direct contact should facilitate communication and remove this roadblock. Similarly, if the specialist is not agressive enough, indirect contact may contribute to a passive role on his or her part. Thus a design team in which all members work in immediate contact is recommended.

Support equipment. Support equipment takes the form of test equipment, special tooling and software packages. Test equipment can be designed and developed within organization and many companies do indeed have test equipment departments. Alternatively, equipment can be bought, but the designer needs to understand the specifications of the equipment and how they relate to the design in question. For example, one does not want to purchase a $30,000 piece of equipment when a $5,000 piece will do.

Special tooling arises when available tools are not adequate to realize the product or manufacture it in quantity. The designer should specify a candidate system which makes use of existing production facilities and often the manufacturability of a product should be taken as a design construct. This requires the designer to have at least some passing reference to production techniques or consult a manufacturing specialist.

Finally, display software packages should be available to allow development of the informational structure. Usually, after development, a program will reside in the display itself in Read Only Memory for special purpose displays.

Prototype Construction

Having evaluated the costs associated with producing the product, a decision is made to initiate a production effort. This requires a prototype to be constructed. However, this construction itself requires the consruction of various subcomponents which must then be combined to yield the overall product.

Design of [illegible]. Parts, according to Asimow (1962), are the elementary building blocks from which components are assembled. For informational structures, parts are labels, framework and specifier. It is through the design of parts that physical realization of the product is acheived.

Design of components. Components are realized from the assembly of parts. They represent formats and organizations of informational structures. As in the case of the parts, layout drawings are necessary to describe the components.

Design of subsystems. The preliminary design stage concerns the overall concept--the candidate system. Subsystems would constitute the informational structure and medium driven by a mainframe, making development more efficient. Subsystems are evaluated in the context of selecting the optimal candidate.

One word of note concerning subsystem optimization is in order. If the constraints of the system as a whole, and those imposed by the requirement of compatibility are not fully recognized and taken into account in the optimation of subsystem, then the final system may be suboptimal with respect to its criterion function. Thus subsystem design must receive system constraints as inputs.

Assembly. After the consituent parts have been prepared, the form of the components can be fixed through final drawings. After the component assemblies are prepared, the subsystems can be similarly fixed. Finally, the system or product as a whole is assembled as a stand alone display.

The computer-aided design and manufacturing (CAD/CAM) technology has contributed to increased productivity, and design cost reductions as well as shortening the product development cycle by automating this design step. Automating this step reduces the manual labor involved in producing drawings, PC boards and schematic diagrams. It also makes redesign economically feasible since updating requires modifying a CRT screen and not redrawing the assembly as a whole.

Test and Evaluation

While the prototype is being constructed the test program can be readied. Tests which verify that the completed product or system will perform in accordance with system requirements are called *verification* tests (Meister, 1971).

The test plan has two major purposes. First, it requires the designer to determine precisely what must be done in the test, thereby avoiding any ambiguities which may later contaminate the data collection. Second, it communicates test objectives and procedures to the client or engineering management. The test phase should contain the following sections (cf. Meister, 1971).

Purpose of Test. General objectives must be broken down into specifics, the purpose of which is to make clear the measures and data collection procedures employed. Thus performance as a whole must be broken down into specific objectives.

Description of system. Since performance data will be collected on specific tasks, performed with specific subsystems and tests using specific procedures, then we must list these.

Test criteria and measures. The measures taken arise from the test objectives and input/output requirements. Subsystem outputs also determine measures. It is also important to note how much data are to be collected. This is usually determined by time and cost considerations. In other cases, statistical reliability may be a consideration.

Data collection methodology. Data collection can take many forms--questionnaires, observational methods, experimental methods, simulation and the like. The particular methods chosen must be listed. Also if any instrumentation is required, it too must be specified. A brief procedural protocol should also be included as well as facilities required.

Sample. We must describe the sampling procedures we will use and the number of units to be tested. If the human is an integral part of the product, then we must specify the number and type of users which will be tested.

Data analysis. We must specify the analytical procedures used to assess the extent of system conformance to performance specification. These analyses are usually constrained by the method used to collect the data.

Test schedule and sequence. The components of the test and the duration of each phase should be listed. Any special reporting procedures should also be specified by the designer.

Redesign

The results of the test and evaluation step may suggest revisions in the design. Revisions are always part of the design process and can occur at every step. Thus, design is an iterative process. If the process set forth is followed and the designer exercises creativity along the way, then the task of redesign will involve primarily minor revisions.

However, as Asimow (1962) notes, all revisions threaten the integrity of the design. The reason for this threat lies in the tightly bound interrelationships existing between the parts of the system. Changing one part may necessitate change elsewhere and so on.

If revisions are needed then, similar to other points in the design process where decisions are necessary, alternative solutions to each should be set forth. This implies a criterion for selecting the most effective solution. This criterion should be based on the integrity of the design. That is, choose the alternative that results in the least amount of secondary changes to the product or system as a whole.

Summary

A process of display design, built around Asimow's (1962) version of the general process, has been presented. The process starts with a feasibility study, the result of which is a set of useful solutions. From this set a 'best' solution must be obtained. Obtaining an optimal candidate sytem is the purpose of the preliminary design activity. This optimal system is then physically realized and tested in the detailed design activity.

It is hoped that the overall strategy for design, together with the framework for display design, will help both the designer or person selecting a display system to make the best decisions possible regarding its implementation.

References

Alger, J.R.M. & Hay, C.O. *Creative Synthesis in Design.* Englewood Clifs, New Jersey: Prentice-Hall, 1964.

Asimow, M. *Fundamentals of Engineering Design.* •lewood Cliffs, NJ: Prentice-Hall, 1962.

Banks, W.W. & Boone, M.P. A method for quantifying control accessibility. *Human Factors*, 1981, *23* (3), pp 299-304.

Bunl, H.R. *Creative Engineering Design.* Ames, Iowa: Iowa State University Press, 1960.

Choudhury, A.A. Multiple criterion interaction evaluations in design. MS Thesis, Department of Industrial Engineering, University of Houston, Huston, Texas, 1979.

Goodman, N. *Language of Art: An Approach to the Theory of Symbols.* Indianapolis, Ind.: Bobbs-Berrill, 1968.

Harris, D. *Fundamentals of Human Factors for Engineering and Design.* Anacapa Sciences, 1982.

Hill, P.H. *The Science of Engineering Design.* New York: Holt, Risehart & Winston, 1970.

Hitt, W.D., Schutz, H.G., Christner, C.A., Ray, H.W. & Coffey, L.J. Development of design criteria for intelligence display formats. *Human Factors*, 1961, 3, 86-92.

Kosslyn, S.M. *Image and Mind.* Cambridge, MA. Harvard University Press, 1980.

Kosslyn, S. & James, P.

Kosslyn, S.M., Pinker, S., Parkin L.P. & Simcox, W.A. *Understanding Charts and Graphs: A Project in Applied Cognitive Science.*

Kreifeldt, J.G. & Hill, P.H. Toward a theory of man-tool system design applications to the consumer product area. Paper presented at the 18th annual meeting of the Human Factors Society, Huntsville, Alabama, October, 1974.

McCormick, E.J. *Human Factors in Engineering and Design.* New York: McGraw-Hill, 1976.

Meister, D. *Human Factors: Theory and Practice.* New York: Wiley, 1971.

MIL-H-46855B Human Engineering Requirement for Military Systems, Equipment and Facilities, 1979.

Ostrofsky, B. Application of a structured decision process for proper inclusion of human resources in the design of a power unit support stand. Air Force Interim Technical Report No. F49620-77-C-0116, 1978.

References (cont'd)

Ostrofsky, B. Design, Planning and Development Methodology. Englewood Cliffs, New Jersey: Prentice Hall, 1977.

Ostrofsky, B., Donaghey, C.E., Marquina, N.E. & Kiessling, E.A. Application of a design morphology to the MX10C[illegible] definition of a fault detection and dispatch system. Air Force Final [illegible]nical Report No. F49620-77-C-0116, 1980.

Pitts, C. Techniques in Engineering Design. New York: Wiley, 1973.

Rogers, D. Display systems development logic. Paper presented at the 25th annual meeting of the Human Factors Society, 1981.

Shurtleff, D.A. How to Make Displays Legible. La Mirada, CA: Human Interface Design, 1980

Simcox, W.A. Cognitive Considerations in display designs. NIE Final Report, 1982.

Van Cott, H.P. & Kinkade, R.G. Human Engineering Guide to Equipment Desig Washington, D.C.: U.S. Government Printing Office, 1972.

List of Figures and Tables

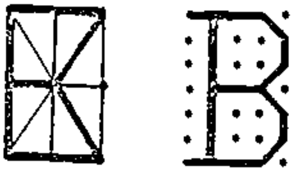

FIGURE 1 Symbols Formed by Fixed-Position and Random-Position Strokes

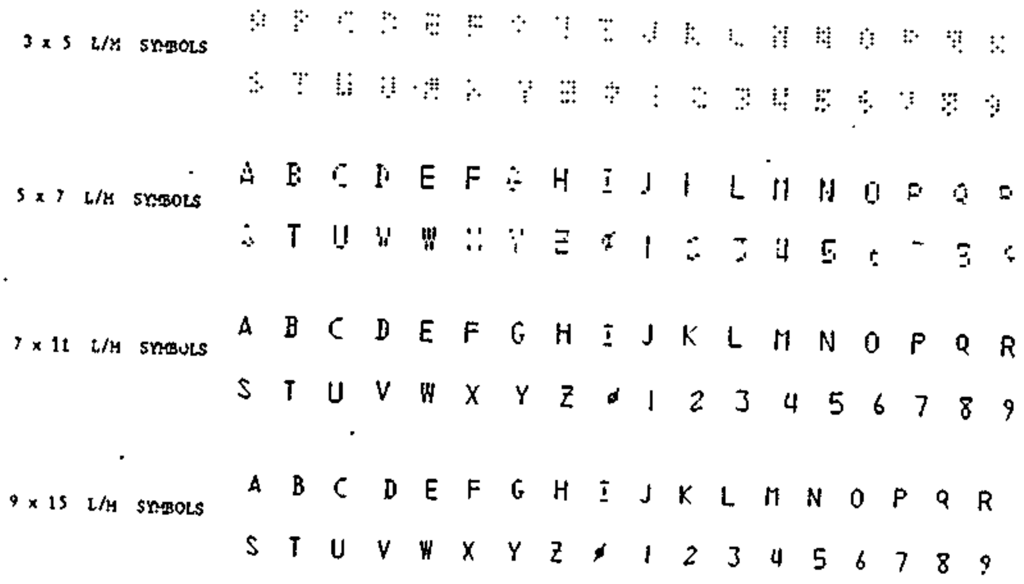

FIGURE 2 Different Size Dot Matrix Symbols

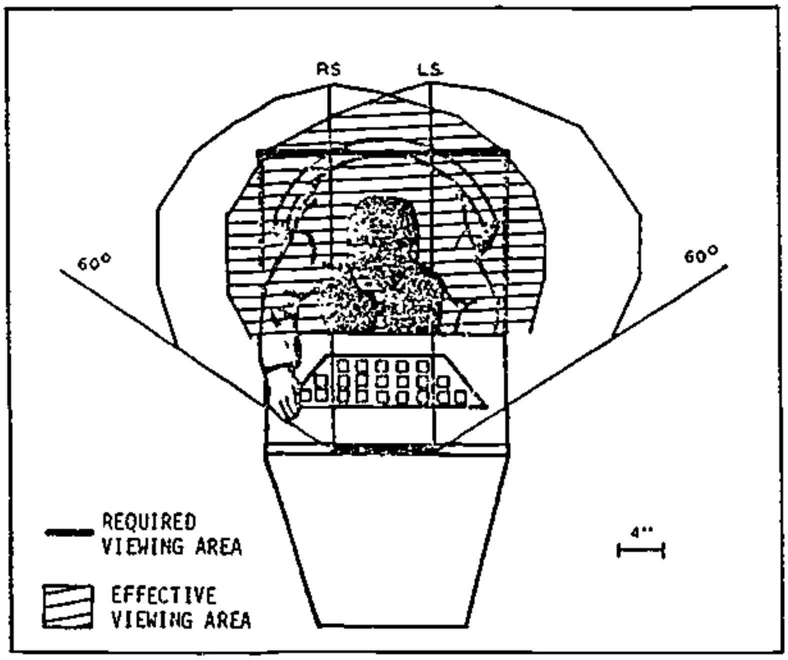

FIGURE 3 Viewing Area Defined From Requirements

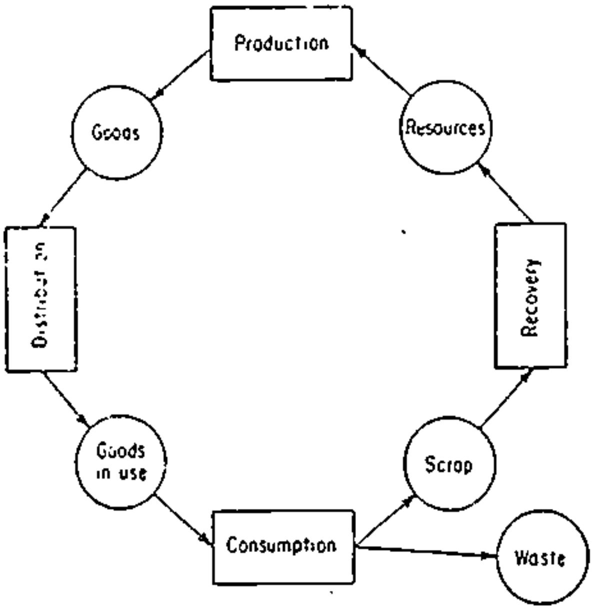

FIGURE 4 Production-Consumption Cycle

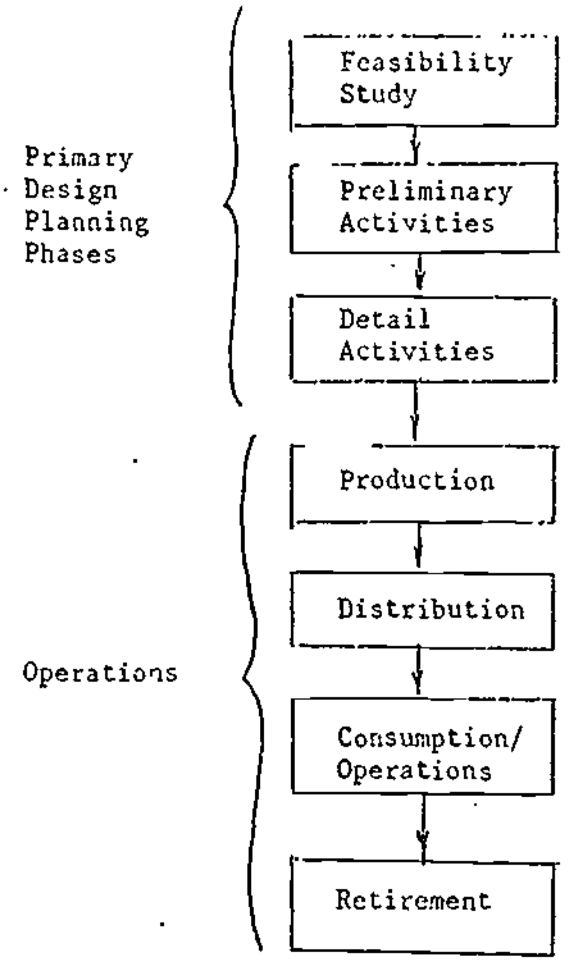

FIGURE 5 Sequence of Design Activities

Feasibility Study

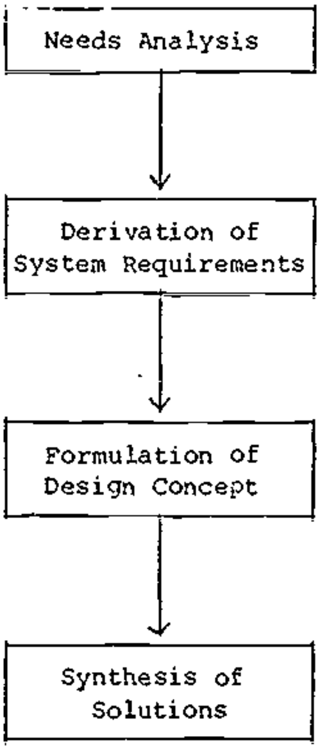

FIGURE 6 Feasibility Study

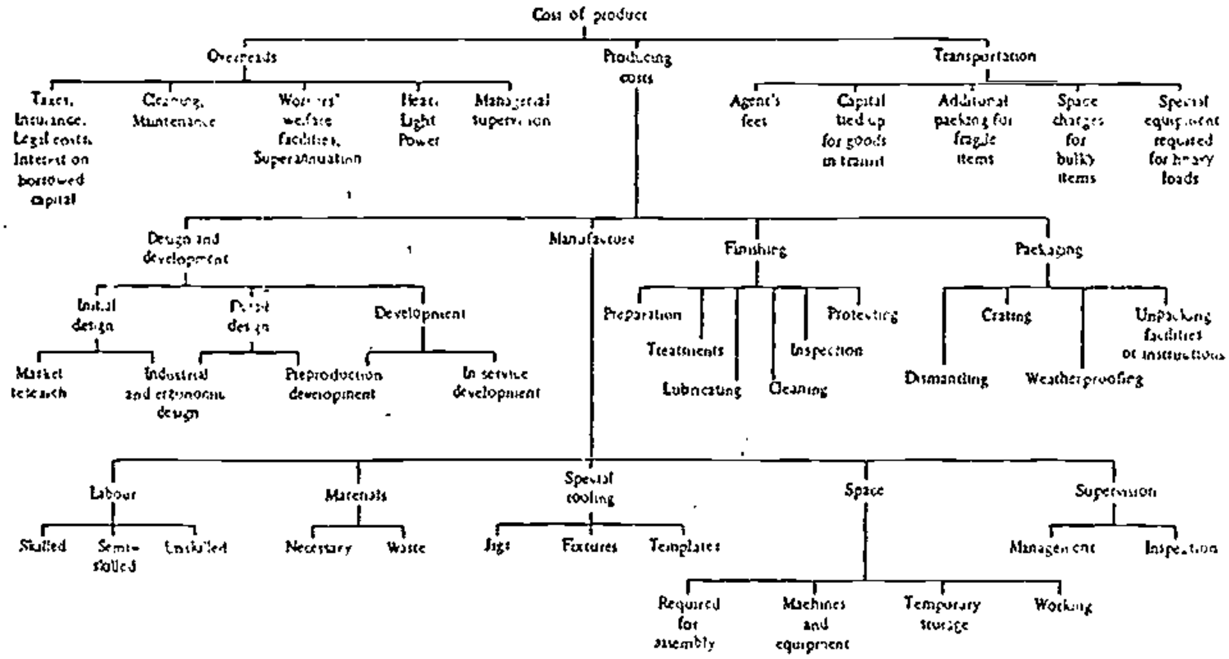

FIGURE 7 Costs Associated with Design

PRODUCTION

INPUT		OUTPUT	
INTENDED	ENVIRONMENT	DESIRED	UNDESIRED
PLANNING AND SCHEDULING MATERIALS MANUFACTURING PROCESSES SAFETY PROVISIONS (i.e. Water, Calcium Hypoclorite)	GOV'T PAPERWORK AEI & SPEC LIST SPO DESIGN REVIEW SPECIAL EQUIPMENT LEAD TIME-SCHEDULE LEAD TIME-PARTS APPROVAL OF SAFETY ENGINEERING	LOW COST SAFE SUFFICIENT CAPACITY GOOD QUALITY FEW COMPLEX ASSYS WELD VS BOLT ASSY DECISION	SPECIAL TOOLING SPECIAL SKILLS SPECIAL EQUIP SPECIAL MATL

FIGURE 8 PRODUCTION PHASE INPUT-OUTPUT MATRIX

OPERATIONS

INPUT		OUTPUT	
INTENDED	ENVIRONMENT	DESIRED	UNDESIRED
WATER SUPPLY TRAINING FOR H-70 HANDLING NITROGEN SUPPLY H-70 SUPPLY TRAIN FOR SYSTEM OPERATION TECH DATA	SERVING INSTRUCTIONS OPERATING RESTRICTIONS 1. VENTILATED AREA 2. RUNNING WATER 3. BREATHING APPARATUS 4. EXISTING SAFETY INSTRUCTIONS (SEE SPEC LIST) 5. GLOVES TIME LIMIT	PROVIDE EPU TANK SERVICING SHOP EASE OF USE: 1. LARGE-EASY TO READ DIALS 2. EASY TO GRASP HANDLES & KNOBS (WHILE WEARING GLOVES) 3. KNOBS AND DIALS, GROUPED TOGETHER WELL LABELED NON-COMPLICATED INSTRUCTIONS	SPECIAL CLOTHING SPECIAL EQUIPMENT BREAK DOWNS HYDRAZINE INJURIES OTHER OPERATING INJURIES

FIGURE [illegible] OPERATIONS PHASE INPUT-OUTPUT MATRIX

DISTRIBUTION

INPUT		OUTPUT	
INTENDED	ENVIRONMENT	DESIRED	UNDESIRED
DEPLOYMENT PLAN INTEGRATED LOGISTICS PLANNING TRANING FOR H-70 HANDLING INTERNATIONAL CONCERNS (EPG ETC) EASY TO HANDLE EASY TO TRANSPORT TRAINING FOR USE AND REPAIR OF EQUIPMENT	ROUGH HANDLING PACKAGING SPECS COMMERCIAL CARRIER	SAFETY IN HANDLING ALL PARTS READY AT SAME TIME UNIT IS COMPLETE	LATE OR UNSATISFATORY EQUIPMENT BULKY-REQUIRING SPECIAL HANDLING

FIGURE 10 DISTRIBUTION PHASE INPUT-OUTPUT MATRIX

RETIREMENT

INPUT		OUTPUT	
INTENDED	ENVIRONMENT	DESIRED	UNDESIRED
USE FOR F-16 PROGRAM OR LIFE OF PRESENT EPU SYSTEM REVIEW POSSIBLE FUTURE APPLICATIONS REPLACE WEAK SUBSYSTEMS	FLUCTUATIONS IN SCRAP MARKET REQUIRED STORAGE FACILITIES WITH PROPER SAFETY FEATURES	SELL FOR SCRAP OR MELT DOWN RECYCLE METAL AND PARTS USE FOR OTHER SYSTEM OR PROGRAM	POLUTION PROBLEMS WITH RETIREMENT

FIGURE 11 RETIREMENT PHASE INPUT-OUTPUT MATRIX

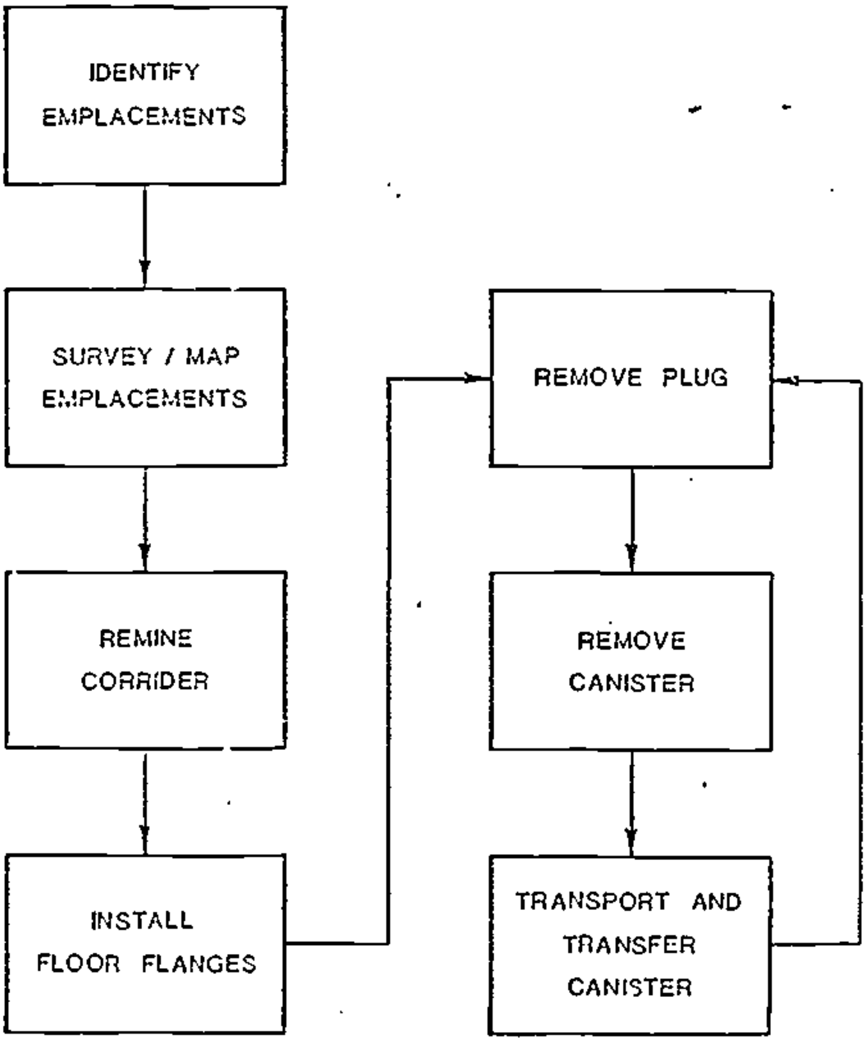

FIGURE 12 Functional Flow Diagram

GROSS SYSTEM FUNCTION: REMOVE CANISTER
ENVIRONMENTAL CONSTRAINTS
Environment is potentially radioactive. Ambient temperature is approximately 100°F. Illumination must be provided. Noise levels could exceed exposure standards. Ventilation must be provided. Humidity levels may be high.
SYSTEM CONSTRAINTS
Removal system must adapt to canisters might not perpendicular to the surface. Operators must be capable of rapid egress. Operator must be kept isolated from canister. Some canisters might be degraded. Emplacement rooms are one-way; work must proceed from back to front.

FIGURE 13 Listing of Functional Constraints

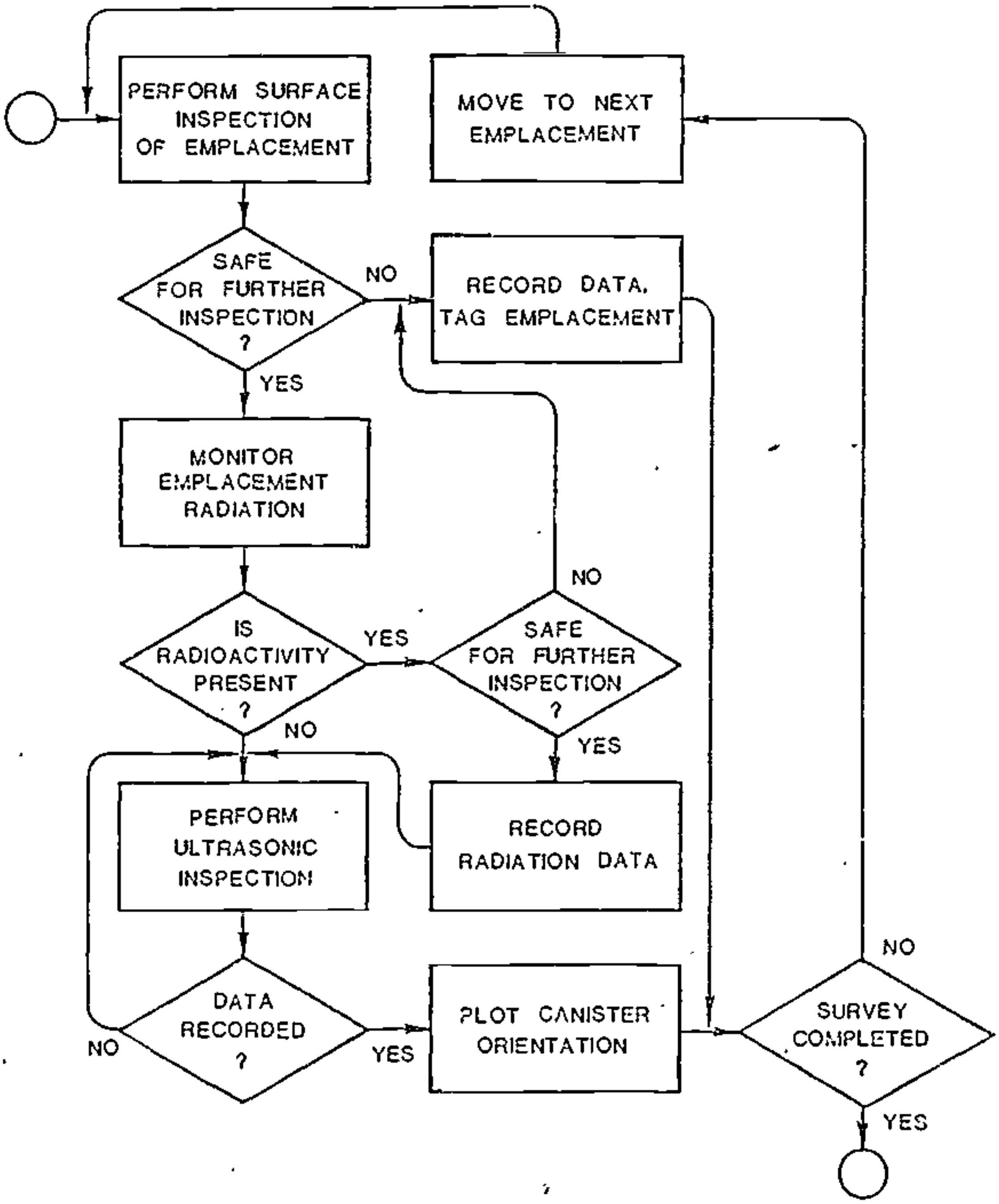

FIGURE 13 Decision/Action Diagram

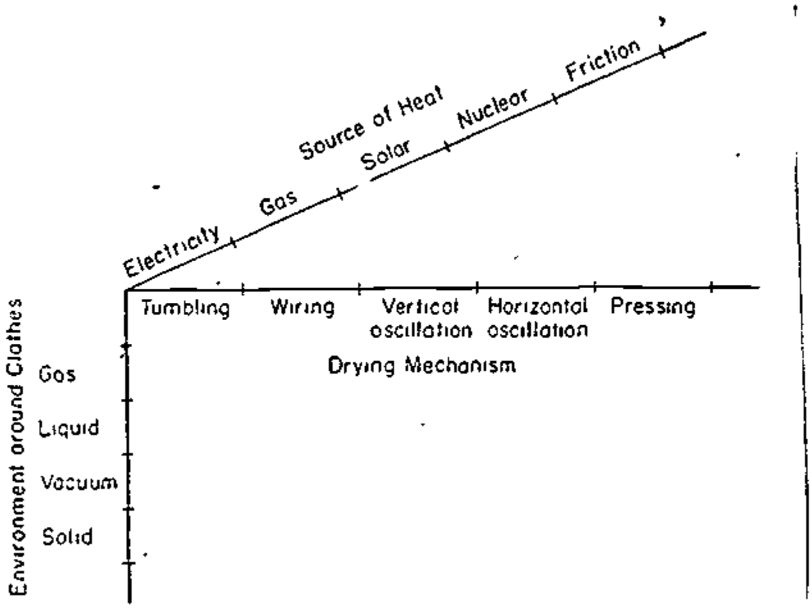

FIGURE 15 Morphological Analysis of Drying Function

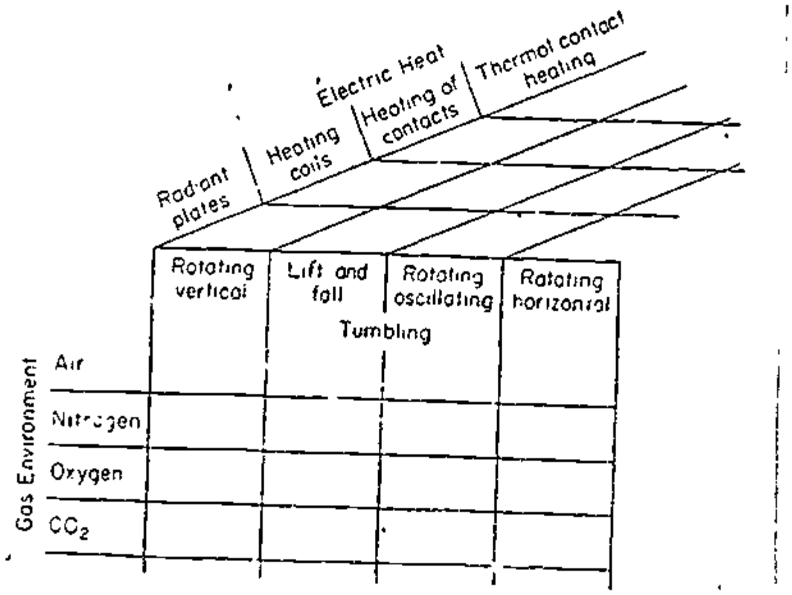

FIGURE 16 Morphological Analysis of a Candidate System

Preliminary Design

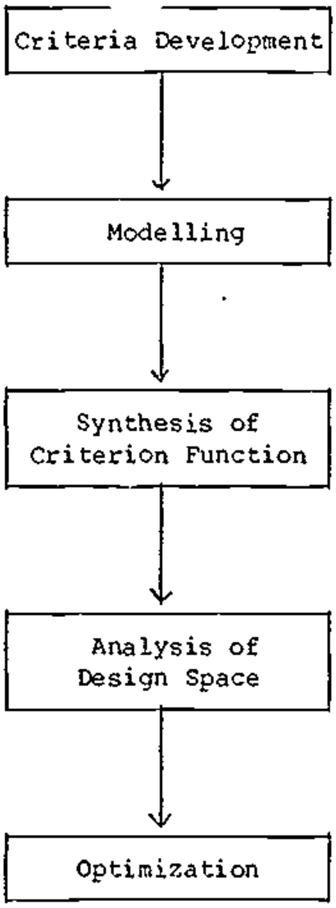

FIGURE 17 Preliminary Design Stage

Detailed Design

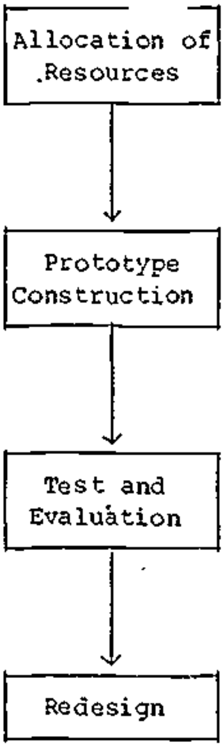

FIGURE 18 Detailed Design Stage

PARAMETER	RECOMMENDED RANGE
Luminance (ft.L)	10 to 50
Luminance Contrast	2:1 to 18:1
Symbol Height (minutes of arc of visual angle subtended at the eye)	10 to 37
Stroke Width (Stroke Width-to-Height Ratio)	1:4 to 1:8
Horizontal Spacing (Percent of Symbol Height)	10 to 65

TABLE 1

Criteria for Display Design - Classical Factors

Graphics Device	Line Quality	Screen Dynamics	Display Capacity	Area Fill	Color Quality
Stroke Writer	Excellent	Excellent	Fair	Poor	Fair
Storage Tube	Excellent	Poor	Excellent	Poor	Poor
Raster Scan	Fair	Fair	Excellent	Excellent	Excellent

TABLE 2

Comparison of Graphic Display Technologies

PARAMETER	RECOMMENDED RANGE
Scan Parameters	
Horizontal Resolution (lines per symbol height)	12 to 18
Bandwidth (MHz)	4 to 20
Direction of Contrast	D/L or L/D
Dot-Matrix Parameters	
Size	5 x 7 to 7 x 11
Spacing of Elements	Contiguous
Stroke-Matrix Parameters	
Method	Random Position
Size	5 x 7 to 9 x 9
Generation Parameter	
Method	Dot or Random Position

TABLE #3

Ranges of Values for Modern Electronic Factors in CRT Design

TABLE Assignment of Weights to Functional Criteria

Criteria	Choice Tally	Total	Weighting Coefficient
1. Performance requirements	1 1 1 1 1 1 1	7	.25
2. Cost	0 1 1 1 1 0 1	5	.178
3. Reliability	0 0 1 1 1 0 1	4	.143
4 Maintainability	0 0 0 1 0 0 1	2	.0715
5 Producibility	0 0 0 0 0 0 0	0	.0
6 Safety	0 0 0 1-1 0 0	2	.0715
7. Number of personnel required	0 1 1 1 1 1 1	6	.214
8 Power requirements	0 0 0 0 1 1 0	2	.0715
		28	.9995

TABLE 5 CRITERIA AND ELEMENTS

SAFETY	CODE
EASE OF MANEUVERING	B
WEIGHT OF FULL TANK	A
VOLUME OF FULL TANK	A
ARRANGEMENT OF CONTROLS	D
PROBABILITY OF LEAKAGE	B
COST	
MEAN COST PER PURCHASED PARTS	A
TIME REQUIRED FOR ASSEMBLY	A
NUMBER OF PURCHASED PARTS	A
OVERHEAD RATE	A
LEARNING CURVE	A
NUMBER OF UNITS BUILT	A
TOTAL NUMBER OF PARTS	A
COST OF MANUFACTURING TIME	A
EASE OF USE	
SIMPLICITY OF PROCEDURES	A
READABILITY OF GAGES	A
SIMPLICITY OF WASTE DISPOSAL TASKS	A
MANHOURS PER SERVICING EPU TANK	A
WEIGHT OF TANK	A
VOLUME OF TANK	A

DURABILITY	CODE
EASE OF SHIPMENTS	A
NUMBER OF OPERATING CYCLES	B
PRODUCABILITY	
TOTAL NUMBER OF PARTS	A
NUMBER OF PURCHASED PARTS	A
AVAILABILITY	
TIME TO MAINTAIN STAND PER DAY	A

A	Directly Measurable
B	Indirectly Measurable
C	Observable

TABLE 6 CRITERIA, SUBMODELS, AND PARAMETERS

	1 SAFETY		2 COST	3 EASE OF USE		4 DURABILITY		5 PRODUCABILITY	6 AVAILABILITY
PARAMETERS Y_K	EASE OF MANEUVERING	PROBILITY OF LEAKAGE		MANHOURS FOR SERVICING TANK	EASE OF MANEUVERING	NUMBER OF OPERATING CYCLES	EASE OF SHIPMENT		TIME TO MAINTAIN STAND
1. NO. OF CONNECTORS		X							
2. WEIGHT OF TANK & H-70	X				X				
3. PRODUCTION MANHOURS PER UNIT			X						
4. MEAN COST PER PURCHASED PART			X						
5. MANHOURS FOR SERVICING EPU TANK				X					
6. SIMPLICITY OF PROCEDURES				X					
7. READABILITY OF GAGES				X					
8. SIMPLICITY OF WASTE DISPOSAL TASK				X					
9. EASE OF SHIPMENT							X		
10. LIFE OF F-16 PROGRAM						X			
11. NO. OF PURCHASED PARTS			X					X	
12. TIME TO MAINTAIN STAND									X
13. TOTAL NO. OF PARTS								X	
14. A/C FLIGHT HOURS PER MONTH					X				
15. NO. OF A/C PER STAND					X				