Section 3

Engineering Design – Process and Principles

3.1 Engineering problem-solving

Engineering is all about solving problems. Engineering design, in particular, is a complex series of events that can involve logic, uncertainty, and paradox, often at the same time. There are a few ‘common-denominator’ observations that can be made about problems in general.

Engineering problems are:

- *Multi-disciplinary* Discipline definitions are largely artificial; there are no discrete boundaries, as such, in the physical world.
- *Nested* Every part of an engineering problem contains, and is contained within, other problems. This is the property of *inter-relatedness*.
- *Interactive* The final solution rarely arrives at once. The solution process is a loop.
- *Full of complexity* So you can’t expect them to be simple.

3.2 Problem types and methodologies

Engineering problems divide into three main types, each with their own characteristics and methodology for finding the best solution. A methodology is a structured way of doing things. It reduces the complexity of a problem to a level you can handle.

3.2.1 Type 1: Linear technical problems

These consist of a basic chain of quantitative technical steps (Fig. 3.1), mainly calculations, supported by robust engineering and physical laws. There is substantial ‘given’ information in a form that can be readily used. Note how the problem-solving process is *linear* – each quantitative step follows on from the last

Engineers’ Data Book, Fourth Edition. Clifford Matthews.

and there are few, if any, iterative or retrospective activities. The solution methodology involves rigorous and accurate use of calculations and theory. Rough approximations and 'order of magnitude' estimates are not good enough.

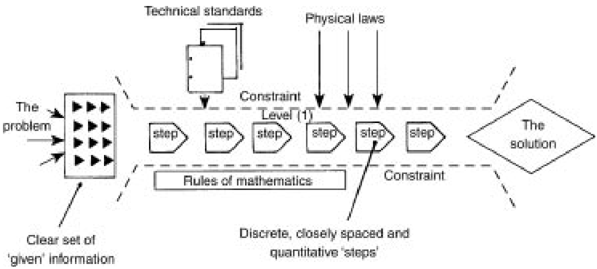


Figure 3.1

3.2.2 Type 2: Linear procedural problems

Their main feature is the existence of procedural constraints controlling what can be done to further define the problem and then solve it. Don't confuse these with administrative constraints; they are established procedural constraints of the *technical* world (Fig. 3.2). The methodology is to use procedural techniques to solve the problem rather than approaching it in an overly technical way. The problem is still in linear form, so you have to work through the steps one-by-one, without being retrospective (or you will lose confidence).

3.2.3 Type 3: Closed problems

These look short and simple but are crammed with hidden complexity. Inside, they consist of a system of both technical mini-problems and awkward procedural constraints (Fig. 3.3).

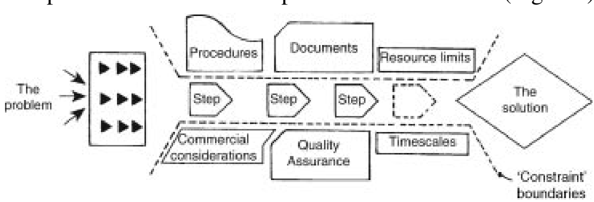


Figure 3.2

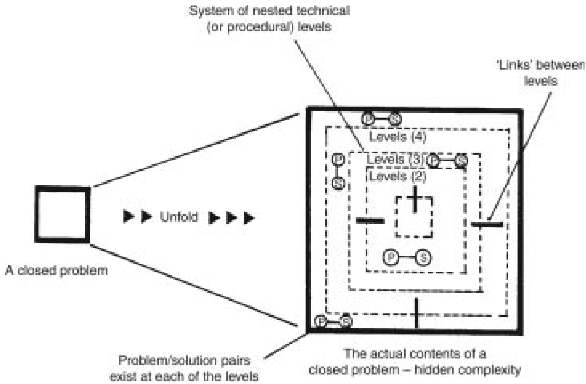


Figure 3.3

The methodology involves 'opening-up' the problem to reveal its complexity before you can solve it. Some hints are:

- Look for *common* nesting levels – you can anticipate these with practice.
- List the *variables and technical parameters* that you feel might be involved – then think for yourself in a pro-active way.
- Think *around* the problem, looking hard for the complexity (you will be revealing it, not introducing it because it is there already).
- Use *group input* – closed problems do not respond well to an individual approach. A group of minds can form a richer picture of a problem than can one.

Remember the golden rule: decide what type of problem you are looking at before you try to solve it.

3.3 Design principles

Engineering design is a complex activity. It is often iterative, involving going back on old ideas until the best solution presents itself. There are, however, five well-proven principles of functional design that should be considered during the design process of any engineering product.

- *Clarity of function* This means that every function in a design should be achieved in a clear and simple way, i.e. without redundant components or excessive complexity.
- *The principle of uniformity* Good functional design encourages uniformity of component sizes and sections. Any variety that is introduced should be there for a *reason*.
- *Short force paths* It is always best to keep force paths short and direct. This reduces bending stresses and saves material. Local closure (in which forces cancel each other out) is also desirable – it reduces the number of 'wasted' components in a design.
- *Least constraint* This is the principle of letting components 'go free' if at all possible. It reduces stresses due to thermal expansions and unavoidable distortions.
- *Use elastic design* Good elastic design avoids 'competition' between rigid components which can cause distortion and stresses. The idea is to allow components to distort in a natural way, if that is their function.

3.4 The engineering design process

The *process* of engineering design is a complex and interrelated set of activities. Much has been written about how the design process works both in theory and in practice.

There is general consensus that:

Design is the use of:

- Scientific principles

  +
- Technical information

  +
- Imagination

Designs are hardly ever permanent. All products around us change – sometimes gradually and sometimes in major noticeable steps – so the design process is also *continuous*. Within these points of general agreement there are various schools of thought on how the process works.

3.5 Design as a systematic activity (the 'pugh' method)

This is a well-developed concept – one which forms the basis of UK degree-level design education. It conceives the process as a basically linear series of steps contained within a total context or framework (see Fig. 3.4).

A central design core consists of the key stages of investigation, generating ideas, synthesis, manufacture, and evaluation. The synthesis stage is important – this is where all the technical facets of the design are brought together and formed into a final product design specification (known as the PDS). The design core is enclosed within a boundary, containing all the other factors and constraints that need to be considered. This is a disciplined and structured approach to the design process. It sees everything as a series of logical steps situated between a beginning and an end.

3.6 The innovation model

In contrast, this approach sees the design process as being circular or cyclic rather than strictly sequential. The process (consisting of basically the same five steps as the 'Pugh' approach) goes round and round, continually refining existing ideas and generating new ones. The activity is, however, innovation-based – it is *creativity* rather than rigour that is the key to the process.

Important elements of the creative process are:

- *Lateral thinking* Conventional judgement is 'put on hold' while creative processes such as brainstorming help to generate new ideas.
- *Using chance* This means using a liberal approach – allowing chance to play its part (X-rays and penicillin were both discovered like this).

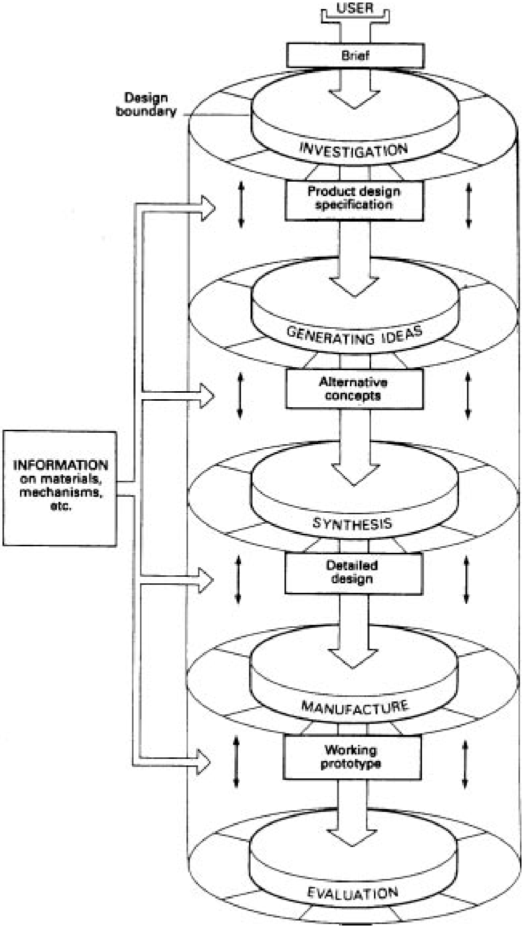


Figure 3.4 Design systemic activity model (overall concept adapted from the model used by SEED in their Curriculum for Design publications)

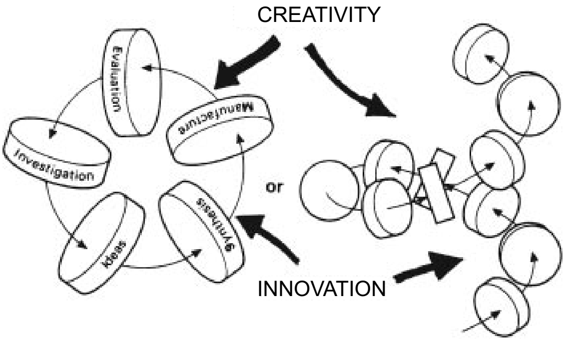


Figure 3.5

- *Analogy* Using analogies can help creativity, particularly in complex technical subjects.

Both approaches contain valid points. They both rely heavily on the availability of good technical information and both are *thorough* processes – looking carefully at the engineering detail of the design produced. Creativity does not have to infer a half-baked idea, or shoddiness.

3.6.1 Design and develop (or not)

It is a strange property of the engineering world that those people or teams that design a product rarely seem very good at developing it. Take any large complex product – a cutting-edge fighter aircraft, an intricately designed medical monitor, or a process flow system of high complexity – and you can see this principle in action. Group necessity spawns the need for the product and individual ingenuity provides the spark that sets the design process in motion – until the combined weight of multiple minds in the design and project team takes over, steamrolling it to its final (hopefully complete) solution.

Now the product is finished, its creators marvel in its complexity, swimming in self-congratulation of the intricacy of its interlocking parts, its form and structure, and the overall purity of its design. The product goes to market, customers are satisfied, and further orders will hopefully follow.

Now the problems start. Within a very short time customers' requirements refine themselves. Paradoxically, the more successful the product, the more extensively it is bought and used, and the quicker the customers' feedback loop works. Almost immediately customers discover nice little improvements that would be desirable – an extra access door here, or a more robust switch there, or a part that is redundant and could be safely omitted. Outside pressures of regulations, design codes and standards change through time, and big customers change their purchasing preferences, or leave the market altogether, to be replaced by others with bigger or more specific ideas about what they want.

Now the competition starts. Successful products breed almost immediate competition. In a well-rehearsed series of events, competitors 'cluster around' the successful product or design, copying its radical ideas, simplifying its design or changing its manufacturing methods to reduce the price. This goes on for a while – the weaker and 'out of their depth' ones soon drop out, leaving a hard core of competitors. And so the market settles.

For the original design team the solution would seem obvious – improve the design using the same initiative, engineering understanding and flair that produced the design in the first place. Surprisingly, this very often proves almost impossible. It is easy to start, and try, but real success is rare. It is as if the technical blinds come down, barriers of various types rise from the twilight of the previous success, and the old flair cannot seem to quite apply itself to improving its own previous creations. There is no simple explanation, but it is a combination of:

- entrenched thinking;
- self-denial that the original can be improved, by anyone;
- belief that the new customers are wrong, and they will eventually realize it and revert to wanting the original product;
- the overall fact that development and iterative improvement of a product is a completely different business to making one

from scratch, best suited to different people with similar, but new, sets of skills.

On a more practical note, the original design team soon move on to other work and projects, leaving a management and budget vacuum that may have form and voice, but little substance. Paradoxically, almost no one can see this. Attempts at development fail after a short time, or drag on interminably, getting nowhere except the regular day-trip to their past glories.

The engineering industry is full to the rafters of examples like this – it seems to be very difficult to continually replicate sparking success. Innovators are best at innovating and are rarely good at development. For those companies and teams that are good at development, they rely (knowingly or unknowingly) on finding a stream of engineering innovation provided by others.

The most complex products, systems or ideas seem to suffer the worst from this design vs development paradox. Whereas the complexity of the original design should, in theory, cement its innovators into first place in their sector of industry, instead it covers them with a cloak of illusion that their advantage will be permanent and that competitors cannot possibly match their ingenuity, coupled with their engineering flair and balance. Time generally proves them wrong.

3.7 Creativity tools

Creativity is important in all facets of engineering design. Many of the developments in creative thinking, however, come from areas outside the engineering field. Figure 3.6 shows the five main creativity tools.

- *Brainstorming* Ideas are put forward by a group of people in a 'freewheeling' manner. Judgement of all ideas is deferred absolutely: no criticism is allowed. This helps stimulate originality.
- *Brainwriting* A version of brainstorming in which people contribute their ideas anonymously on slips of paper or worksheets – these are then exchanged and people develop (again anonymously) each other's ideas in novel ways.

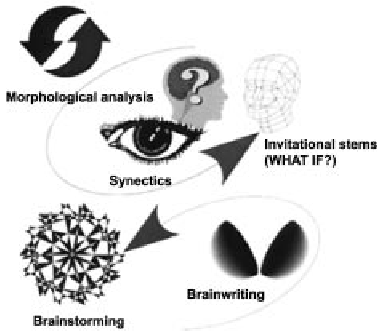


Figure 3.6

- *Synectics* A specialized technique including the joining together of existing, apparently unrelated, ideas to reveal new perspectives or solutions to a design problem.
- *Morphological analysis* This is a formal, structured method of solving design problems using matrix analysis.
- *Invitational stems (wishful thinking)* A loose and open creative process encouraged by asking questions such as 'wouldn't it be nice if' or 'what if material cost wasn't a problem here?'.

3.7.1 Useful references

A key introductory paper to the subject is:

Thompson, G. and **Lordan, M.** 1999, A review of creativity principles applied to engineering design, *Proc. Instn Mech. Engrs*, *Part E*, *J. Process Mechanical Engineering*, **213** (E1), pp. 17–31. This paper includes a list of over 70 detailed reference sources.

A list of publications concerning creative design methodologies can also be found at: http://www.isd.uni.stuttgart.de/~rudolph/engdesign_publications.html.

3.8 The product design specification (PDS)

Whatever form the design process takes, it ends with a PDS. This sets out broad design parameters for the designed product and sits one step 'above' the detailed engineering specification.

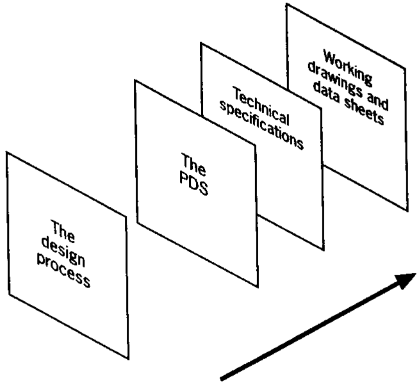


Figure 3.7

The Product Design Specification (PDS) checklist

- Quantity
- Product life-span
- Materials
- Ergonomics
- Standardization
- Aesthetics/finish
- Service life
- Performance
- Product cost
- Production timescale
- Customer preferences
- Manufacture process
- Size
- Disposal
- Market constraints
- Weight
- Maintenance
- Packing and shipping
- Quality
- Reliability
- Patents
- Safety
- Test requirements
- Colour
- Assembly
- Trade marks
- Value analysis
- Competing products
- Environmental factors
- Corrosion
- Noise levels
- Documentation
- Balance and inertia
- Storage

3.9 Presenting technical information

3.9.1 Technical information – what is it?

Technical information is information that has its roots in some sort of technique or method. It can be theoretical, practical, or a subtle mixture of the two and can be thought of as a specific form of language – the language of technology and industry. A further common factor is that technical information is related to the application of a technological skill, either in producing the information itself or using the messages it conveys.

Where is it used?

Your will see technical information used:

- in all fields of pure science;
- in all the applied sciences;
- as the bedrock of all the technical disciplines and subjects that you can think of.

Because of its wide application, the variety of types of technical information is itself wide and complex. Some disciplines (computer technology, for example) have almost a separate technical language of their own, but the majority of technical disciplines thrive on forms of technical information that have fairly general application.

What is it for?

Technical information conveys ideas between people. This is an important point – despite the proliferation of computer-generated data of all types, the prime purpose of technical information is to convey ideas, concepts, and opinions about technical matters between people. These may be rough ideas, elusive and fleeting concepts, or finely honed technical proofs and axioms – all come under the umbrella of technical information.

Does it have any other uses?

Yes. Technical information is a tool of persuasion, and the way it is presented plays a part in convincing people of others' understanding and opinions about technical subjects. All scientific and technological activities hinge around the way that technical ideas are transferred between the participants – it is this flow of technical ideas that gives a technology or a project its direction. This means that you can think of the presentation of technical information (in all its forms, remember) as perhaps *the* most common tool of science and technology.

Presenting technical information – the challenge

The world of technical information is beset with the problem of complexity. The rich technical variety that exists in every technological discipline manifests itself as an ever-increasing amount of complex information that has to be presented in an easily digestible form. The task of presenting technical information is, therefore, about finding *simple ways* to present *complex ideas*. In most cases, algebraic or mathematical expressions become too complicated to be understood by anyone but academics, so it is better to find other ways. Five guidelines are presented below.

Some guidelines

- Use graphical methods of communication wherever possible.
- Supplement algebraic and mathematical information with geometry to make it simpler and/or clearer.
- Use visual *models* to portray ideas.
- Don't be afraid of making approximations where necessary.
- Use sketches, diagrams and drawings.

There is one common factor in these five points – they all involve the use of *models* to present technical information effectively. The task of presenting technical information is therefore about constructing a representation of that information, so that its meaning can be conveyed on a computer screen or printed page.

The need for imagination

Many of the skills of effective technical presentation involve the use of imagination. Although traditional methods are well established there is always room for improvement and adaptation. Trends over the past ten years have favoured the increased use of graphical and pictorial information in preference to tables of mathematical and algebraic data – such modern presentation methods *need* the use of imagination to keep the development and improvement going. Imagination is also needed in the choice of method to be used for the presentation. It is difficult to keep technical presentations looking fresh and interesting if you use the same technique too often; you have to look for alternative ways to convey your information.

Making the choice

For any situation in which you have the task of presenting technical information, you are faced with several general choices:

- Tabulation (i.e. lists of tables of data)
- Graphical methods
- Scientific or symbolic representation
- Technical drawings of some sort
- Pictorial representation, such as sketches and three-dimensional diagrams

The choice between these is best helped along by learning to do a bit of critical thinking. Ask yourself a few questions about the situation, such as:

- Which method will help me to present this technical information in the *clearest* way?
- Is this method really suitable for this type of information, or am I just using it for convenience?
- Does this method have visual power — or does it look mediocre?
- What are the positive and negative aspects of the method I am about to use?

Remember that the purpose of this type of critical thinking is to help you to choose a good presentation technique, not to stifle any imagination you are trying to bring to the process. Viewed from this perspective, the task of presenting technical information begins to resemble a process of technical problem-solving – a logical choice between alternatives, coupled with a bit of imagination and flair to liven up the result.

Presenting technical information: A summary

- *The purpose* of presenting technical information is to convey technical ideas, facts and opinions between people. It is also a tool of persuasion.
- There are always several different ways to present any set of technical information.
- The challenge is to find simple ways to present complex ideas. This leads to five main principles of information presentation:
  - Graphical methods
  - Combining information with diagrams
  - Using 'models'
  - Making approximations
  - Using sketches and diagrams (of various types)

After all of this you have to use a little imagination and flair — and then make a decision about the best presentation method to use.

3.9.2 Categories of information

The way to better understand the general subject of technical information is to think of it as being divided into wide but precise categories. An understanding of the existence of these categories will also help you to think critically about the purpose of different types of technical information, and how to present them in the best possible way. Figure 3.8 shows the situation. Note the three main categories: guidance only, symbolic/schematic, and prescriptive – with all three capable of belonging (at the same time) to categories of information that can be described as being *in*ductive or *de*ductive. Now look at the categories in turn, as shown in Fig. 3.8.

Guidance-only information

Not all technical information is presented in a form that provides an exact description of something (an object, procedure, or idea). Its purpose is merely to give you guidance – to convey a

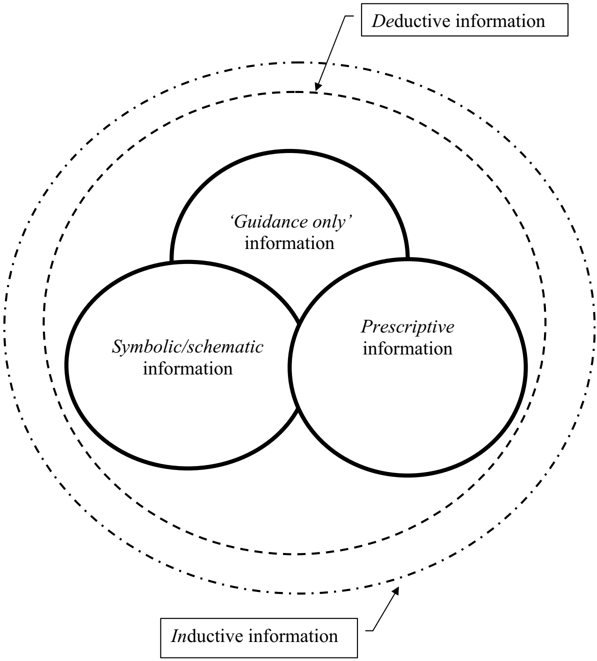


Figure 3.8 Different 'categories' of technical information

general technical idea. To do this, the method of presentation often involves approximations about:

- Fundamental relationships between, for example, technical procedures, designs or physical objects.
- Trends in size or movement.
- The physical shape and layout of objects or components.
- Dimensions.

The best way to understand this is by looking at two examples. Figure 3.9 shows a representation of a simple

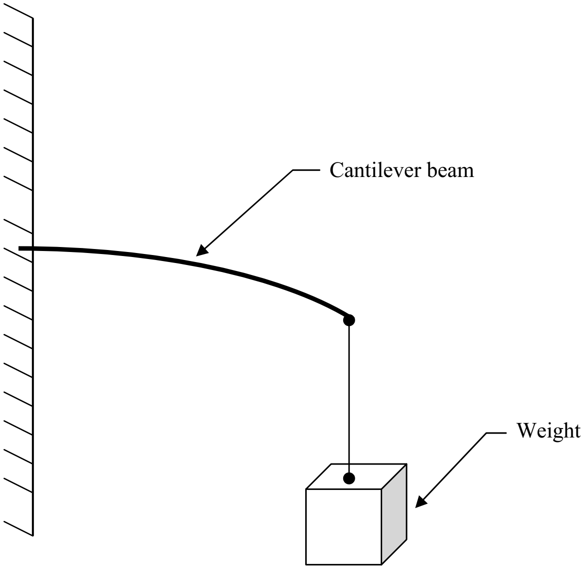


Figure 3.9 An example of 'guidance only' information

cantilever beam supporting a suspended weight. This is a graphical presentation, because it is in the form of a picture. From a quick study of the figure you should be able to infer three main points:

- The diagram is intended to show you that the beam *bends* under the influence of the weight, but not precisely how far it is bending, or the exact shape of its curvature.
- The diagram is showing you a 'general case', which would be applicable to all cantilever beams, because there is no attempt to show the length or cross-section of the beam, its material, or the even the size of the weight on the end.
- The beam and the weight, and the physical way that they relate to each other, are represented by a drawing that has a strong resemblance to the real visual world, i.e. it is similar to

the way the objects look in real life. There are obviously some approximations – the beam itself, for example, is represented by a thin line with no apparent depth or thickness, and there are no proper mechanical details of how the end of the beam locates into the wall, or how the weight is fixed to the free end. None of these, however, would stop you recognizing the physical arrangement of such a loaded beam if you saw one, so the drawing is a close *representation* of the real-life object rather than being merely a symbol.

In summary, the message that this diagram provides – i.e. that cantilever beams bend when a weight is applied – is really of use as 'guidance only'. It is a non-precise but important part of the technical picture – not exactly correct, but good enough. This type of information can be useful in many areas of technology, particularly in engineering design disciplines where technical ideas are developed in a series of steps.

Symbolic/schematic information

Symbolic and schematic types of information are so closely related that they are best thought of as a single category.

Symbols and schematics

What is a symbol?

A symbol is something that represents something else by association, resemblance, or convention.

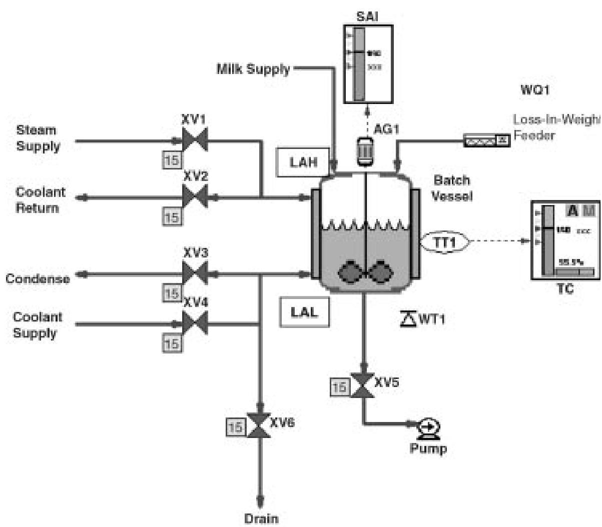
What is a schematic?

A schematic shows the scheme or arrangement of things, normally by using symbols to artificially reduce complexity.

In practice, many methods of presenting technical information are a merger of the symbolic and schematic approaches. This is valuable in just about all technical disciplines as a way of simplifying a complex system, object, or set of technical relationships down to a level that a reader can understand. In many cases, symbolic/schematic representations are the only way to portray complex technical information in a user-friendly form. Typical examples are:

- Process Instrumentation Diagrams (PIDs) for any type of process plant.
- Hydraulic, pneumatic, electrical and similar circuit diagrams.
- Applications where it is necessary to show the *structure* of something or how it works (such as Fig. 3.10).
- Symbolic illustrations (see Fig. 3.11) which portray technical information and look nice.

One common thread running through schematic representations is that they show directly, or infer, physical interrelationships between parts of things, often in the form of a schematic plan or design. In contrast, pure symbolic representation (as in Fig. 3.11)



A process system 'schematic'

Figure 3.10 An example of schematic information

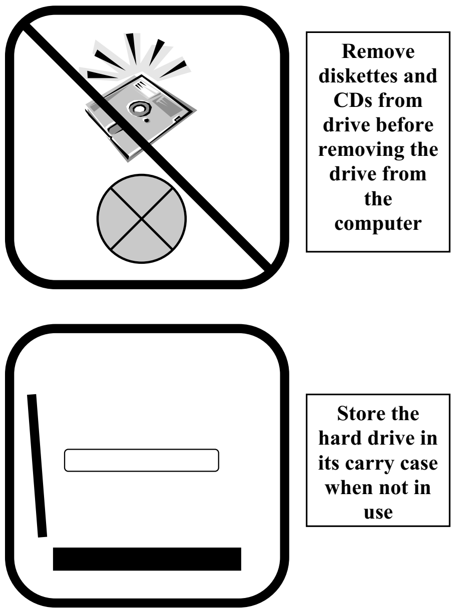


Figure 3.11 An example of symbolic information

can be more 'stand alone', or simply give a small piece of technical advice.

Prescriptive information

'Prescriptive information' is information that sets down firm rules, or provides an exact description of something. Note that the term contains the noun *prescript*, meaning a direction or decree. Not surprisingly, technical information that is prescriptive is generally complex, because it is not always possible to fully describe complex things in a simplified or shortened way. It can also have an air of rigidity about it, rooted in the fact that it is attempting to explain the unique and detailed solution to a difficult technical situation or problem.

The power of prescriptive information lies in its ability to cause people to take action, evaluate a system in a particular way, or assemble a series of engineering components in the correct order. Prescriptive information is often seen (and used) 'nearer' the single solution of technical problems – in contrast to guidance information, which is more of an upstream technique that is used during earlier stages where the technical atmosphere is more diverse and conceptual.

Prescriptive information
You can expect to see prescriptive information:

- in precise mathematical routines and algorithms;
- in manufacturing procedures;
- in instruction manuals; and
- in any situation where technical information contributes to step-by-step problem-solving.

A further feature of prescriptive information is its accuracy. Unless technical information is accurate in number and expression (as in mathematical or algebraic notation) and in form (i.e. shapes or spatial representation) it cannot really be prescriptive, because it would leave too much freedom, thereby hindering the achievement of a unique pattern or solution. This is why prescriptive information is particularly suited to technical and engineering disciplines – it thrives on hard-edged ideas.

Deductive versus inductive information
You can think of deductive and inductive information as features of the technical background against which various presentation techniques are applied, rather than discrete presentation mechanisms in themselves. A particular set of technical information may be predominantly inductive, deductive, or (more likely) a subtle combination of the two, with at least part of the definition coming from an understanding of how that technical information was *derived* rather than its effectiveness in conveying technical ideas. In short, this means that you only

need to consider the deductive vs inductive qualities when you come to 'fine-tune' presented information. You don't need it in the earlier stages.

Deductive versus inductive information

Deductive information has a clear link between some previous statement (called the *premise*) and the deduced information (or *conclusion*) that is presented. If the premise is true then it is deduced that the conclusion must also be true. Compare this to the inductive situation where the premise may give support to the conclusion but does not guarantee it. Common examples are:

- Mathematical and algebraic expressions: i.e. $x + x = 2x$. Here x is the premise of sorts, and $2x$ is the conclusion, obtained when x is added to another x.
- Engineering drawings are primarily deductive because they describe (and so rely on) tightly controlled physical relationships between mechanical components that are determined before the drawing is produced.

Inductive information is information that infers a future conclusion based on previous (historical) information or happenings. Examples are:

- Statistical Process Control (SPC) in manufacturing, where the characteristics of components which are not yet manufactured are inferred by previous observation of similar already-completed parts.
- Most empirical laws (e.g. in fluids or mechanics) in which we draw conclusions about a large group of things from observations of one or two specific cases.

You should now be able to see how the differences between inductive and deductive information can be built into the way that technical information is presented. Technical theories, alternatives, and concepts are suited to the use of inductive information because it is never intended that the information is absolutely traceable to a proven premise. Think how this applies to chemistry, materials science, and the gas laws. Newton's laws of motion are also empirical (they have no proof as such) so dynamics and kinematics are disciplines in which technical information is presented in an inductive form. Once these routines are applied for use in engineering disciplines, however, and are metamorphosed into engineering *designs*, the technical information becomes heavily *de*ductive, as it takes its place in the search for precise solutions to engineering problems. Remember that these definitions only become relevant when presentation methods for technical information become heavily refined – they have little relevance to the simpler, (often 'guidance only') forms of presentation.

Graphical methods

The term 'graphical' refers more to the way that information is presented than the nature or purpose of the information itself. The character of some types of technical information makes it particularly suitable to being displayed in graphical form. Graphs are best at showing *relationships*. These may be tight algebraic relationships linked by rigid constants and coefficients, or softer more inferred ones providing information in the form of general guidance and trends. The power of graphical methods lies in their ability to provide answers to several questions at once. A single graph can, if correctly constructed, hold information about:

- linear and non-linear relationships;
- equalities;
- inequalities;
- relationships in time and/or space;
- looser concepts such as regression, correlation and trend.

Because of the complexity that can result, graphical presentations need to be properly ordered if they are to communicate their information clearly.

The character of graphical presentation allows for a wide variety of different types, but brings with it the corresponding disadvantage of an equally wide variety of distortions and misinterpretations. The effective visual impact of graphs means that it is easy to show information in a way that is capable of misinterpretation. You can also make it persuasive or misleading, if that is your intention.

Conventions

Conventions play a pivotal role in the presentation of technical information. They are used in both algebraic and graphical methods to bring uniformity to the way that information is presented (and interpreted) whilst still allowing a degree of flexibility. Don't confuse this with a set of rigid rules, which also bring uniformity, but at the expense of variety and imagination. The conventions themselves are simple – you can think of them as lowest common denominators of the presentation techniques. Some conventions are detailed below.

Scalar methods—Scalar methods use quantities (*scalar* quantities) that have a single 'real number' dimension only — normally magnitude or size. So, any presentation technique that compares information on size only can be loosely referred to as a scalar technique. Figure 3.12 gives some examples. Scalars have the following advantages:

- they are simple;
- quantities can be added, subtracted, and compared using algebraic methods such as addition and subtraction.

Vector methods—Vector methods have more than one dimension; normally size and direction. The appearance of this second quantity is important, as it creates the conditions for illustrating multiple types of information about the subject being presented. Figure 3.13 shows some examples. Vector methods are:

- detailed (or can be);
- useful for showing complex technical situations in many different technical disciplines;
- more difficult to compare with other forms of information — you will not always be comparing 'like with like'.

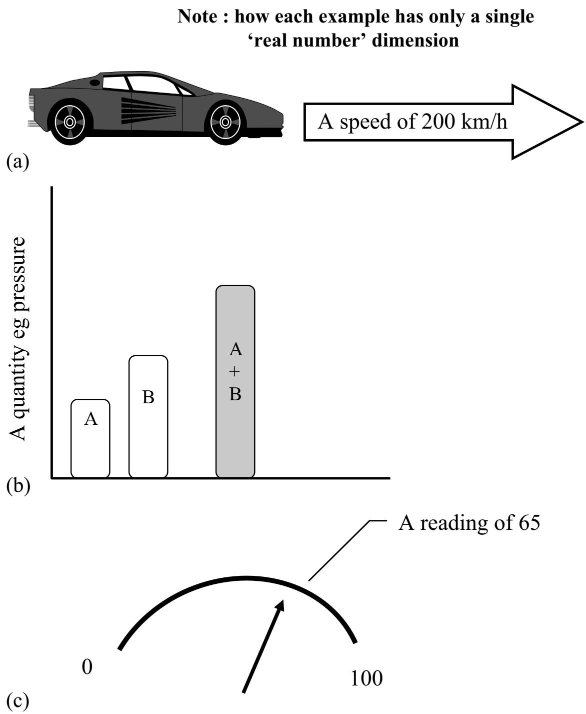


Figure 3.12 Examples of scalar methods

***Matrix methods*—**A matrix is simply a particular type of framework in which information is contained. A common use is as a way to represent a system of mathematical equations containing several unknowns. This forms part of the subject of matrix theory which, together with linear algebra, is used to present information and solve problems in disciplines such as pure mathematics, analysis of structures, thermodynamics and fluid mechanics. Such matrices take the form of an array of 'elements' enclosed in brackets (Fig 3.14(a)).

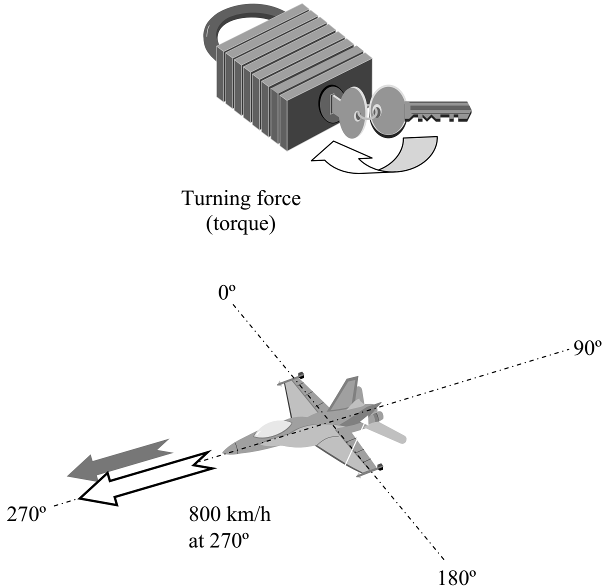


Figure 3.13 Examples of vector methods

$$M_1 = \begin{pmatrix} -2 & 0 & 1 \\ 15 & 3 & -1 \\ 0 & 1 & 2 \end{pmatrix}$$

Figures enclosed in brackets or vertical lines

$$M_2 = \begin{pmatrix} -2 & 0 & 1 \\ -4 & 3 & 8 \\ 0 & 1 & 2 \end{pmatrix}$$

Each matrix *row* represents one equation - the numbers are the coefficients of the variables in the equations

Figure 3.14(a) A set of mathematical *quantitative* matrices

Matrices are also used in their more general sense to display technical information that can be contained in an arrangement of rows and columns. They can exhibit qualitative data about things that have multiple properties, and are particularly useful for use as a selection tool in the design process. Figure 3.14(b) shows a typical example.

This matrix shows qualitative data about design options

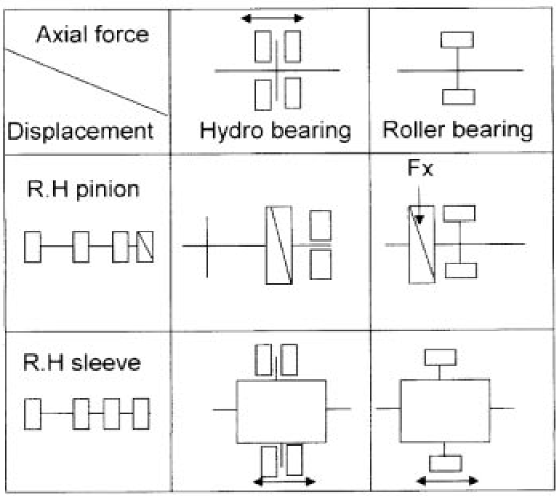


Figure 3.14(b) A *qualitative* use of a matrix

***Dimensions*—**It is convention that graphical and pictorial information can be presented in either one-dimensional (1-D), two-dimensional (2-D), or three-dimensional (3-D) form.

1-D forms, such as simple line graphs, often look as if they are in 2-D format but in reality only convey a single 'dimension field' of information (see Fig. 3.15(a)) – which could if necessary be conveyed by single lines. 1-D information is therefore, by definition, capable of being conveyed by the use of simple lines of negligible thickness (Fig. 3.15(b)).

2-D information conveys information relevant to either two spatial dimensions (x and y axis for example) or to two alternative 'dimension fields' (see Fig. 3.15(c)). As most diagrammatic and pictorial information is presented in 2-D format, it has wide application across the technical and engineering disciplines. 2-D presentations are also useful in that they can masquerade as 3-D views in applications such as wireframe drawings.

3-D presentations are used to portray pictorial views of technical objects. There are several types, each with their own advantages and disadvantages.

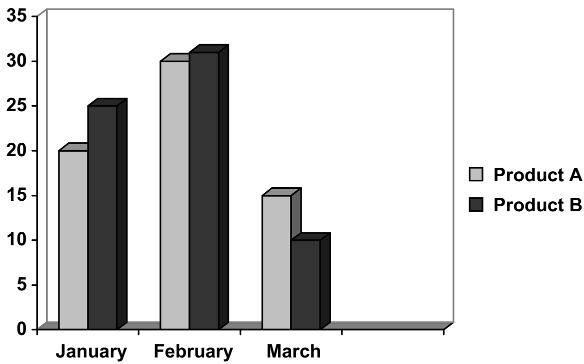


Figure 3.15(a) Sample of a 2-D or 3-D information

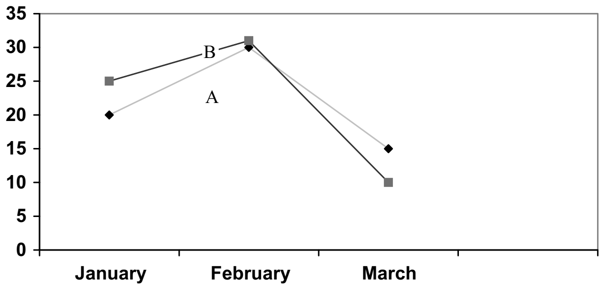


Figure 3.15(b) The actual 1-D message of Figure 3.15(a)

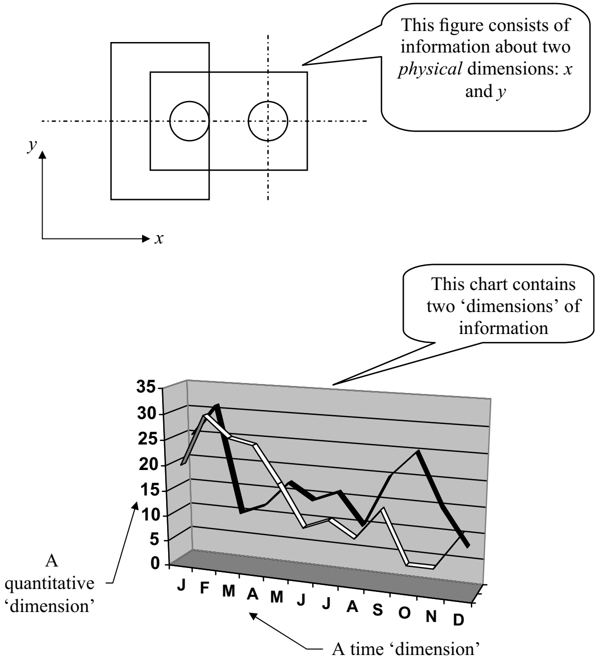


Figure 3.15(c). Two different types of 'two-dimensional' information for Figure 3.15(a)

Reminder – Conventions

Conventions act as the unwritten rules of technical presentation. They bring a level of uniformity to the way that information is shown. The main ones are:

- scalar vs vector presentation;
- matrix conventions;
- 1-D, 2-D and 3-D methods.

Remember that these conventions apply to all forms of technical presentations, not just simple ones in which the conventions may be instantly apparent.

***Co-ordinates*—**Co-ordinates are a method used to locate the position of points, lines and objects in space. They are relevant to most forms of technical presentation that involve accurate graphs or drawings. Figure 3.16 shows the two main co-ordinate systems — note that they can be expressed in either 2-D or 3-D form. The Cartesian system using x, y, z axes and their positive/

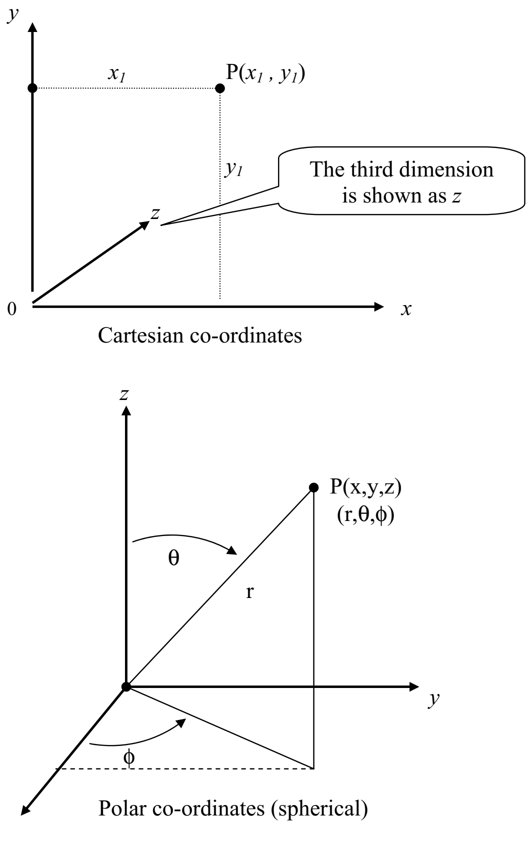


Figure 3.16 The main co-ordinate systems

negative sign conventions is more commonly used for 3-D application than the 'polar' system, which is easier to depict when limited to use in two dimensions. Readings in either of these two systems of co-ordinates can be easily converted to the other. Note, however, that the fundamental differences in their sign convention means that they are rarely seen being used together.

Presenting technical information – Key point summary

- Technical information has its roots in some sort of technique or method.
- It is used in all technical subjects (see Fig. 3.17)
- Good presentation of technical information involves understanding the traditional methods and then applying some imagination.
- Technical presentation is about *choice*. There are often several different ways to show the same thing.
- The important categories of technical information are:
  - Guidance-only
  - Symbolic/schematic
  - Prescriptive
  - Deductive and inductive
  - Graphical methods
- These categories are often cross-linked and combined.

Don't forget conventions such as scalar, vector and matrix methods, and the two main co-ordinate systems: Cartesian and polar.

3.10 The anatomy of mechanical design

Stripped bare, mechanical design is a bit like economics; driven by one main thing: s*carcity*. Scarcity is the state when something is in short supply. There is not as much to go round as everyone would like, so the desirability goes up. In stark contrast to economics, however, engineering design is

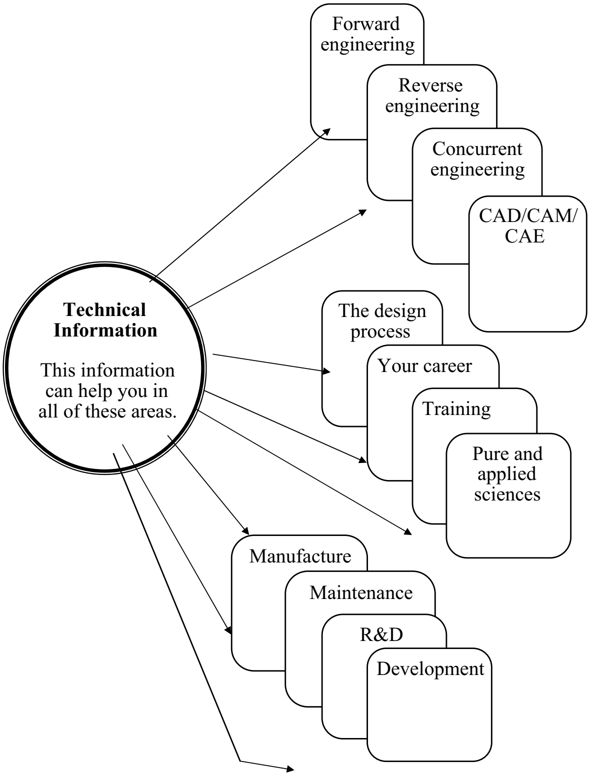


Figure 3.17 Using this information

composed of things you can see, feel and measure, rather than loose concepts and theories surrounded by elusive woolly clouds of this and that.

How does this scarcity show itself? It starts as the driving force that produces the *need* for an engineering product or design. Think of the scarcity of coal leading to the drive to design the steam engine to pump out the mines that produced the coal, and you've got the idea. In a similar way, the scarcity of specialist computer programming skills led to the drive to produce the graphical-user interface with its easy-to-use on-screen icons.

Scarcity really comes into its own, however, when you consider its effect on the *product* (design) side of the equation. The final design of almost any mechanical product is the end result of a series of competitions for things that would make the design 'better', but are in short supply. This is a change from thinking of the design process as a set of linked activities, expressed perhaps in some kind of schematic flowchart or diagram, but it is a fairly accurate reflection of what actually happens.

Figure 3.18 shows the five main things that suffer from the eternal problem of scarcity. Try as they might, there is no way they can escape from it. It is always there, an unseen guiding hand, controlling how each feature can grow and flourish, and how they relate to each other. It's a competition, remember.

The fight for space

In any design, physical space is always at a premium. Small, lights cars are great, but if you need space for multiple occupants they get bigger, and so heavier. Similarly, if you need space for luggage you have to achieve it at the expense of reduced passenger-seating space. The allocation of space to competing design functions (e.g. space for luggage vs space for people) is loosely called *disposition*. Modern design is getting very good at optimizing this. Think of these examples.

- *On-screen keyboards for compact computers*. You need a minimum size of screen to be able to see it, so why not use it to show the keyboard, rather than competing for space with a separate one?
- *Flexible seating in MPV vehicles*. Folding, sliding and/or removable seats allow valuable space to be configured for occupants ***or*** luggage, depending on the need.
- *3-D modelling*. Computer-package modelling of pipework, etc., layouts enables complex piping systems to be shoe-

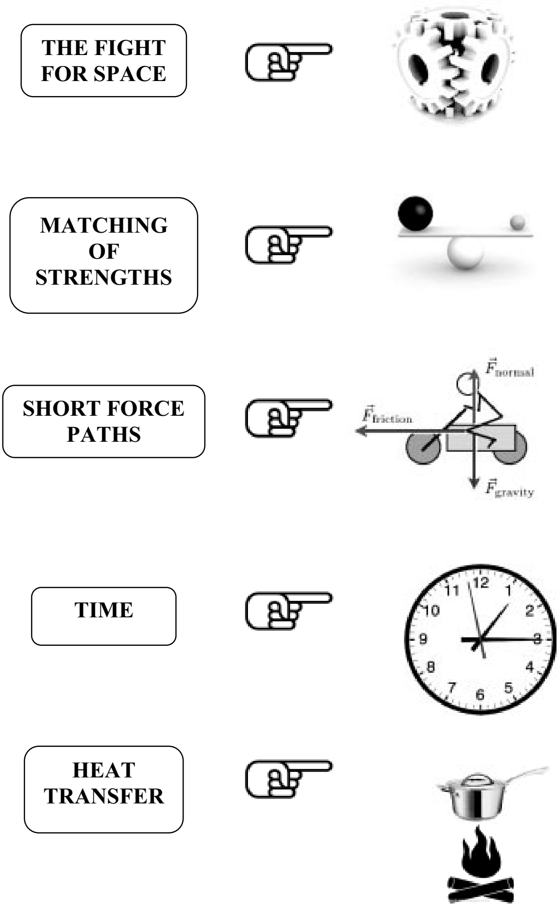


Figure 3.18 Five design factors that suffer from scarcity

horned into the smallest possible space, reducing pipe lengths, weight and therefore cost.

The matching of strengths

Any chain is only as strong as its weakest link. It makes no sense to have oversized components, consuming space and cost, connected to undersized weak ones, which will break before the strong component has even had half a chance to impress its neighbours. For structures, efficient design is about optimizing the scarce resource of the capability to resist load. Each member (strut, tie or beam) would ideally be stressed to the same percentage of the allowable stress it can safely withstand. As load (and, if applicable, time) increases, then eventually all the components would break, neatly and on cue, at exactly the same time.

For mechanisms (containing dynamic elements that move), a similar situation applies, but this time it is not just about force paths, but the transmission of movement, work and power. Think of torsion (a common method of transmitting power, as in an engine or gearbox) as simply a specific case of component loading. The idealized objective is once again to load each member to an identical degree, keeping the sizes of all interconnecting shafts, gears and similar components to the ideal level when, with an excessive increase in load, they would all fail together.

Short force paths

For structures and mechanisms one of the keys to both strength and space issues is that of using *short force paths*. Force (static) and movement (power transmission) both need a path to follow in order that they can appear at a location other than the point at which they were applied. Figure 3.19 shows the idea. Good designs have force paths that are kept short. This keeps bending and torsion to a minimum. Conversely, long force paths encourage bending, resulting in large deflections or torsional distortion. These are particularly bad ideas when acceleration is involved, which will exaggerate the effect of forces.

Examples of short force paths are hidden away in their hundreds in everyday engineering objects. Figure 3.20 shows another example.

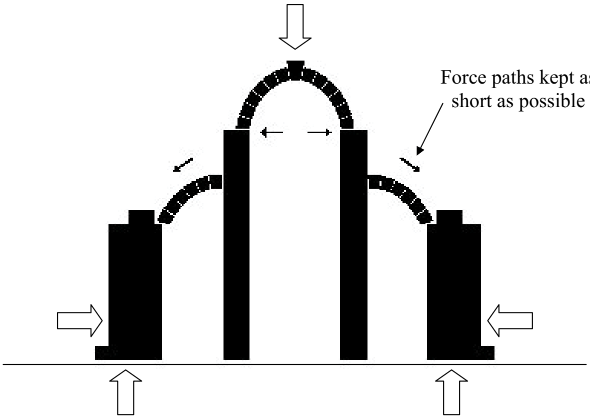


Figure 3.19 Force paths; The flying buttress

Time waits for no one

For mechanical items containing moving parts, as opposed to static structures, time is an important constraint. On balance, it is keeping time to a *minimum* that is the most difficult, i.e. it is relatively easy to get something to happen slowly, but much more difficult to get some movement or action completed quickly.

Physical movement infers a mechanism of some sort. Components such as engines and gearboxes contain hundreds of individual machine elements involved in a variety of rotating, reciprocating, twisting or bending motions. All these movements must happen in the correct place, and at the right time to make the component work as intended.

Conceptually (and practically), the control of time is easy. The time that a component will take to travel from A to B is governed by only two things:

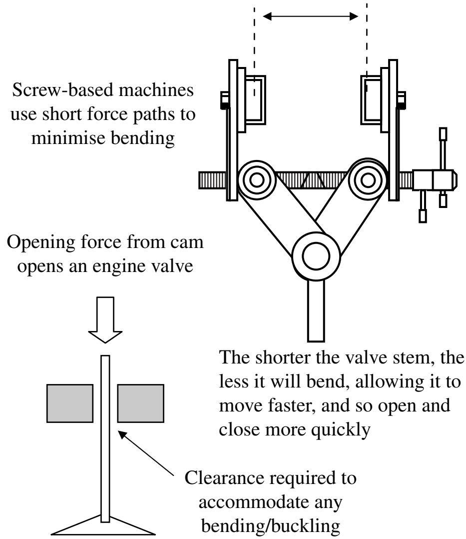


Figure 3.20 Examples of short force paths

- its velocity between A and B; and
- the physical distance from A to B.

If the design need is to reduce the time, this can be achieved by either making the velocity higher, or the distance shorter, or a little of both. Reducing times allows you to make:

- engines faster
- switches more positive
- pumps more efficient
- almost any mechanical component better.

Here are some common examples:

- *Cycle 'Derailleur' gears* (Fig. 3.21). The chain jumps between sprockets with a very quick, positive movement when changing gears. This innovative design goes back to its incarnation in the 1930s but still remains the best way to do the job.

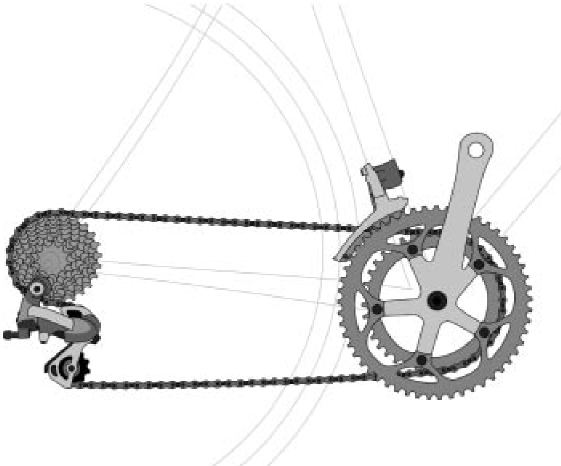


Figure 3.21 Derailleur gear; an example of a 'quick change' mechanism

- *Diesel engine injectors*. Fuel injection into a diesel engine cylinder needs to start, and stop, very quickly. Rapid electronic triggering, with very small physical movement of the injector components, allows the injector to open quickly, and then shut quickly (strong springs give a high velocity). This allows a higher rpm, bringing the potential for increased power.

Good design is therefore about short movements, high speeds and minimal clearances. Any clearance required between two mechanical components (to allow for thermal expansion, misalignment, inaccurate manufacture or whatever) represents 'dead motion', i.e. motion (and hence time) that is wasted. Tight clearances of course require a high accuracy of manufacture.

If the distance between A and B cannot be reduced, the only other solution is to increase the velocity of the component, so that it gets from A to B more quickly. Sadly, constant velocity is a rare concept. Many engineering mechanisms and rotational machines, such as turbines, have constant velocity in some components, until they are subject to vibration when it all goes horribly wrong owing to reciprocating movements. For a component to reciprocate it has to stop and start at either end of a period of movement. This applies whether the reciprocation is planned and predictable (like simple harmonic motion, for example), or random and non-sinusoidal. Stopping and starting brings *acceleration* and its opposite and equally-damaging cousin *deceleration*. On balance it is these that cause components to break rather than velocity. Therefore, in the search for keeping movement time to a minimum, acceleration is a major enemy.

The race from hot to cold

Engines of all types, and many other machines, rely on thermodynamics to make them work. The laws of thermodynamics, all discovered hundreds of years ago, sit quietly in the wings controlling precisely what you can and cannot do. Their advantages allow you to design machines around thermodynamic cycles that work and their limitations tell you precisely what you *can't do*. Try to circumvent these, and you will fail, every time.

There is little more to thermodynamics than the fact that heat will move from hot to cold. A hot body will always lose heat to a nearby cold one, if it can find a route to travel by. Change the speeds of this transfer at various times and places and you have the building blocks of a thermodynamic cycle: adiabatic, isothermal, Carnot or whatever. Once you find a way to get something useful out of the cycle (heat from a boiler cycle, power from an engine cycle, or cooling from a refrigeration cycle) then all you have to do is build the physical parts around it, and away you go (Fig. 3.22 shows an example).

Getting heat from hot to cold is easy because physics will do it for you. Thermal design, therefore, comes down to managing the *transfer* from hot to cold. It is not so much a race as a carefully choreographed procession, with all the bits of heat moving at the correct speed and at the right time. As with the

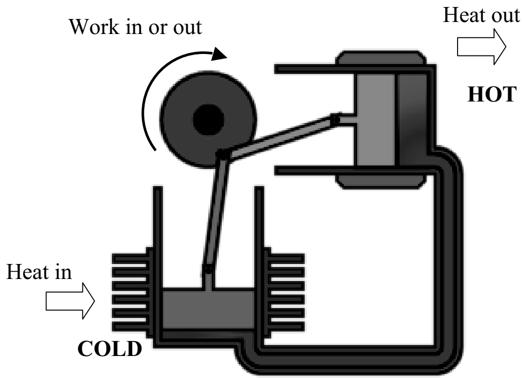


Figure 3.22 Engine cycles: managing the transfer from hot to cold

previous parameters we saw of strength and space, it inevitably ends up as a compromise. The difficulty comes once again in controlling the timescale of the whole affair. Remember the three possible methods of this transfer:

- Conduction
- Convection
- Radiation.

Conduction is fairly predictable. As the conductivity of most materials is well known, the design choice is so difficult. The problem tends to come from combining conductivity with the mechanical properties required. High conductivity materials can be quite weak (copper and aluminium are good examples) so expect the answer to be a compromise.

Radiation is no great mystery either. Properties such as surface emissivity and reflectivity are well defined. Physical distance and space arise again, however, as constraints. On balance, convection is the most difficult to predict, relying on empirical (experimental) results to validate the output of software models. Engine cooling systems, boilers and refrigeration systems are examples of convective systems that developed gradually by iteration and trial and error, rather than appearing in final form almost overnight.

And finally cost

Remember that all of this design activity costs money, and a continuous stream of expenditure can be expected from start to finish.

It is easy to think of cost as a separate category in the anatomy of mechanical design, but of course it is not. It is the direct function of all the others put together. Think of it as a 100% unbreakable correlation – the longer and deeper, and the more iterative the design process, the more that it will all cost. This takes us full circle to the start – the force of scarcity that is the overall driver of the design process. Remember our definition of scarcity:

- *Scarcity is the state when something is in short supply.*

We then have a match where money is used as the universal indication of scarcity. If something is in short supply, the cost will always rise, although the time it takes to do so will vary. Under this set of game rules, the dangers become obvious: designs are driven by the need to reduce cost; products made on a budget may not do their job well or may soon break down; and the marketing and branding departments will then try to make the product look good by presenting and packaging it as a quality item. And there you have it – on the shelf.

Summary

Design is driven by Scarcity

Make sure you use the engineering manifestation of this to work for you because, if you don't . . .

The only better thing about *your* product is that it will cost less.

3.11 Safety in design – principles and practice

Straight thinking about safety in design involves linking together the principles and the practice; both have an engineering basis tempered by common sense.

3.11.1 Safety principles

There are three fundamental principles of design safety, as shown in Fig. 3.23. These are, in *ascending* order of effectiveness: warnings, protection, and avoidance.

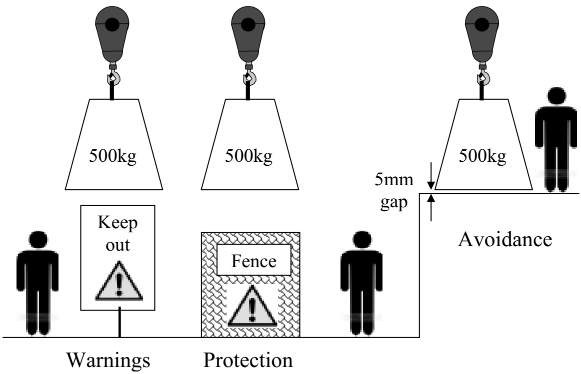


Figure 3.23 Principles of design safety. Reprinted from *Case Studies in Engineering Design*, Clifford Matthews, © Elsevier 1998

Warnings

Warnings are the least effective method of trying to ensure design safety, but still the most common. Whilst signs and instructions may warn of a danger, they don't make it go away, or make its effect any less. Warnings are passive and don't work particularly well.

Protective features

This is an indirect method of design safety. It comes in many forms but all have the characteristics of relying on *protecting* people from a danger (or failure) or mitigating its effects in some way. The danger itself, however, does not go away. Some typical examples are:

- ***Control and regulation systems*** **–** These give a basic form of protection by keeping the operation of a process or item of equipment within a set of predetermined safe limits. This minimizes the danger of failure. An example is the combustion control system on a boiler which acts as a safety system because it limits, albeit indirectly, the steam pressure.

- ***Design diversity*** **–** This is a more subtle principle. The idea of using different design principles within the same design gives protection against common-mode faults and failures. It almost guarantees that any fault that does occur unexpectedly will not be replicated throughout the entire design – so it will not 'fail' completely. Computer software systems are a classic example of this approach, but it is also employed in mechanical equipment by using different materials, types of bearings and sealing arrangements, fluid flowpaths, speeds, etc.
- ***Factors of safety*** **–** Factors of safety, design margins, conservatism, prudence – these are all protective instruments, used to reduce the risk of dangers caused by failure. They apply to mechanical, electrical and electronic systems. Sometimes they are carefully calculated using known properties and failure modes, but at other times they are chosen as a *substitute* for detailed design knowledge. In all too many engineering designs, data on material performance or the state of loading of individual components, and a unanimous understanding of how things actually fail, are imperfect. Fatigue life calculations are a good example in which factors of safety are more *palliative* than prescriptive.
- ***General protective devices*** **–** Again, there are different types: safety valves on boilers, overspeed trips on engines; overcurrent devices – fuses and reverse power protection on electrical equipment being typical examples. Larger equipment installations have separate protective systems, often with features such as duplication, diversity and a self-monitoring capability, to keep the level of risk down. A related type of indirect design safety feature is the *lock-out*, which prevents a component or piece of equipment that is in a dangerous state from being put into operation.

Avoidance features

Avoidance is the principle of achieving safety by choosing a design solution that eliminates danger from the outset. It is by far the most direct (and best) method of ensuring design safety – although not always possible. Eliminating potential dangers starts at the embodiment stage of the engineering design process and feeds forward into the engineering specifications for a piece of equipment. There is often a link, of sorts, with some of the protective features described previously. You can think of danger avoidance features as fitting neatly into three separate principles: fail-safe, safe-life, and redundancy.

The fail-safe principle

This is not quite the same as having protective devices. The idea is that a piece of equipment is designed to *allow for* a failure during its service life but the design is such that the failure has no grave effects. The failure is *controlled*. There are a few ways to do this:

- First and foremost, there must be some way of identifying that the failure has happened – it must be signalled.
- The failure must be *restricted*, i.e. for a machine, it must keep operating, albeit in a limited or restricted way, until it can be safely taken out of operation without causing danger.
- The implications of failure of a single component, need to be understood and assessable as to the effect that it will have on the total machine or system design.

All these presuppose that the consequences of failure are properly understood by the designer – a precise understanding of the definitions is less important than the need to develop a clear view of how a component can fail and what the consequences will be. A useful tool to help with this is the technique of *fault tree analysis* (FTA). You may see variations of this, a common one is 'failure modes effect analysis' (FMEA). There are small differences between them, but the principles are the same. To perform an FTA, you list all the possible modes of failure of a design, and display the consequences of each failure in a network or 'tree' diagram. Figure 3.24 shows an example for a bolted steam pipe joint. Note how the tree starts from the smallest, most divisible components and moves 'outwards' to encompass the design of the 'system' (the bolted

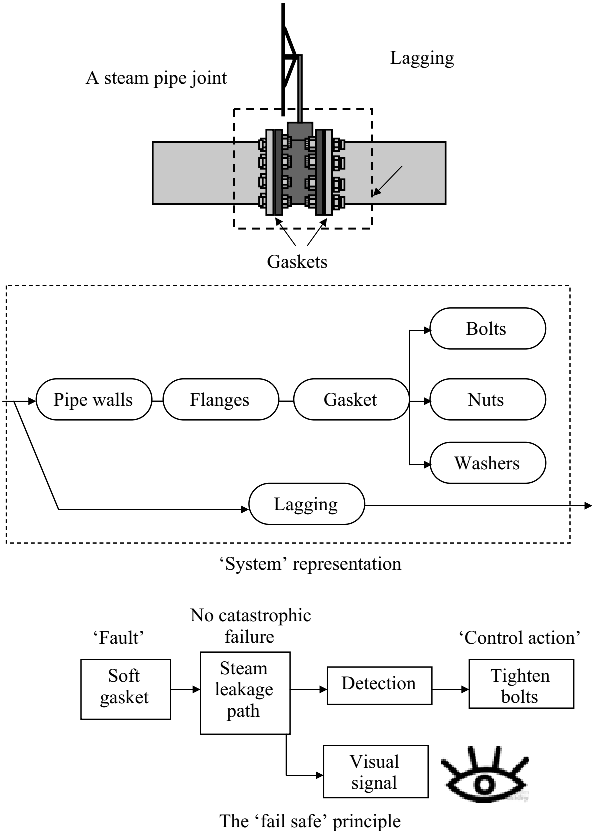


Figure 3.24 A typical component system analysis. Reprinted from *Case Studies in Engineering Design*, Clifford Matthews, © Elsevier 1998

joint). FTA is of greatest use on complex designs consisting of interlinked and nested systems. Oil and gas installations, nuclear plants, airliners, weapons systems and most electronic products are assessed using FTA techniques.

The safe-life principle

This sounds rather obvious: all design components (mechanical, electrical or electronic remember) need to be designed for an adequate *lifetime* to ensure that they won't fail during their working life. In practice, it is not as easy as it sounds – safety is based exclusively on accurate quantitative and qualitative knowledge (yours) of all the influences at work on a component. It is, frankly, almost impossible to do this from scratch every time you do a design – you have to rely on previous, proven practices. There are four areas to consider:

- Safe *embodiment* design based on proven principles and calculations. This means technical standards and codes of practice —but note that not all of them contain embodiment design details.
- Careful specification of *operating conditions*. Operating conditions for engineering components have to be described fully. Fatigue, creep and corrosive conditions are important for mechanical components as they have a significant effect on material lifetime. For electrical equipment, environmental conditions (heat, dust, dampness, sunlight, etc.) can soon reduce lifetime.
- Safe operating *limits* – –again, it is easy to overlook some of the operating limits of engineering components. Low-cycle fatigue and high-temperature creep cause the most mechanical failures if they are overlooked at the design stage. Stresses due to dynamic and shock loadings are another problem area. These failures often occur well within the estimated lifetime of a component.
- Analysis of *overload conditions*. It is not good enough just to consider normal working stresses, currents or speeds, you need to look for the overload condition.

Many mechanical equipment designs have an accepted way of calculating their projected lifetime. Contact bearings are designed using well-proven lifetime projections expressed as an 'L-number'. High-pressure boilers and steam vessels' technical standards specify calculation methods for creep and fatigue life. Safety-critical items such as structures for nuclear reactors, aircraft, tall buildings and high-integrity rotating plant are also designed in this way. There are, however, numerous items of equipment for which the technical standards do not address lifetime. The common mechanical engineering standards covering, for example, steel castings and forgings place great emphasis on specifying detailed mechanical and chemical properties, but hardly mention fatigue life. To compensate, manufacturers of specialized forged and cast components do in-house tests and develop their own rules and practices for defining (and improving) component lifetime.

The redundancy principle

Redundancy is a common way of improving both the safety and reliability of a design – it is also easily misunderstood. The most common misconception is that incorporating redundancy always increases design safety but there are many cases where this is not true. What do you think of the following statement?

- *An airliner with four engines is safer than one with two engines.*

There appears to be some logic in this. On long-haul Atlantic routes the theory is that a four-engined aircraft can suffer two engine failures and still complete the journey safely using the two remaining engines. This is fine for some types of engine failure, but what if an engine suffers a major blade breakage and the damaged pieces smash through the engine casing into the wing fuel tanks? Here the redundancy has no positive effect, in fact there is a counter-argument that four engines have twice the chance of going wrong than do two. The central message is that redundancy is not a substitute for the proper design use of the fail-safe and safe-life principles. Redundancy *can* increase design safety, but only if the redundant components are themselves designed using fail-safe and safe-life considerations. Some specific examples of design redundancy are shown in Fig. 3.25. Note the different types, and the definitions that go with them. These definitions are not rigid, or unique – their main purpose is to help you to think about and identify the different options that are available.

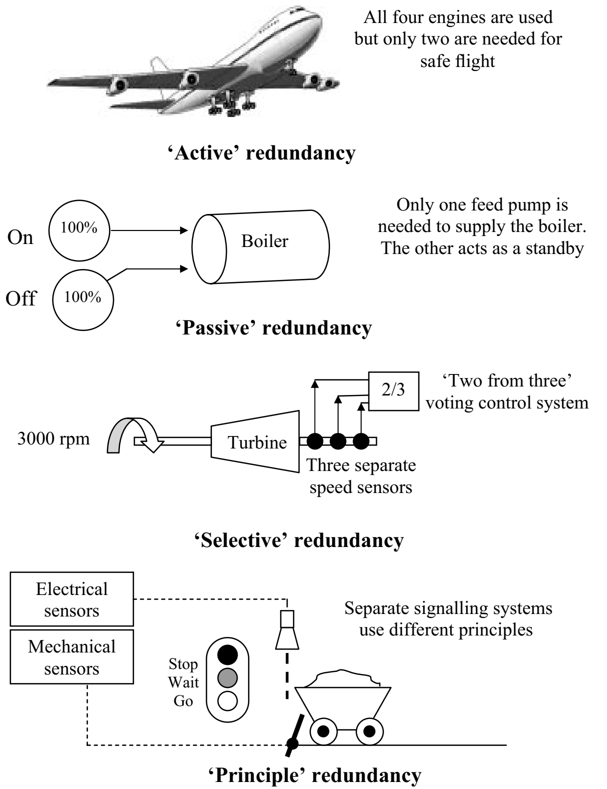


Figure 3.25 Types of design redundancy. Reprinted from *Case Studies in Engineering Design*, Clifford Matthews, © Elsevier 1998

3.11.2 Embodiment design safety

Embodiment – A re-think

Embodiment design is what happens in the rather large grey area between conceptual design, and detailed engineering design where a system or machine is described by a set of specifications and drawings. Embodiment is therefore about *deciding engineering*

The Danger	Comment
Stored energy	Potential and kinetic energy can be dangerous. Stored pressure energy is a particular hazard if released in an uncontrolled way.
Rotating machinery	Rotating belts, couplings gears, fans — anywhere where there is relative movement between a machine and humans is a potential danger.
'Crushing and trap' gaps	Gaps of more than about 8 mm between moving parts can trap fingers.
Exposed electrics	Exposed electrical equipment is an obvious danger.
Hot parts	Components or fluids above about 50ºC will cause burns.
Falling and slipping places	If there is a place possible to fall or slip, someone will always find it.
Noise	Excessive noise is a recognised industrial hazard

Figure 3.26 Typical design 'dangers'. Reprinted from *Case Studies in Engineering Design*, Clifford Matthews, © Elsevier 1998

features. Hence, deciding safety features is part of the embodiment process but is not all of it; other engineering considerations (the main one being *function*) have also to be included. The process of identifying general design safety features is easier than choosing between all the available alternatives. This is because, unlike, for example, the mechanical strength of a component design, safety cannot easily be expressed in quantitative terms, so you often have to work without clear-cut acceptance criteria that you can use to compare the safety level of different designs. It is easy to make general statements about design safety, but not so easy to translate these into the language and features of embodiment design. The best place to start is with a list of design *dangers* and then consider them as you think, in turn, about each part of a design. These design dangers are different, in detail, for each individual component or system design, depending upon what it does, but the general principles are common. Figure 3.26 shows a typical list of design dangers. We can use these as part of a series of steps to tease out good safety features during the embodiment process.

Step 1: Split the design into*systems*

Any technical design, simple or complex, can be thought of as consisting of interconnected systems. These systems come

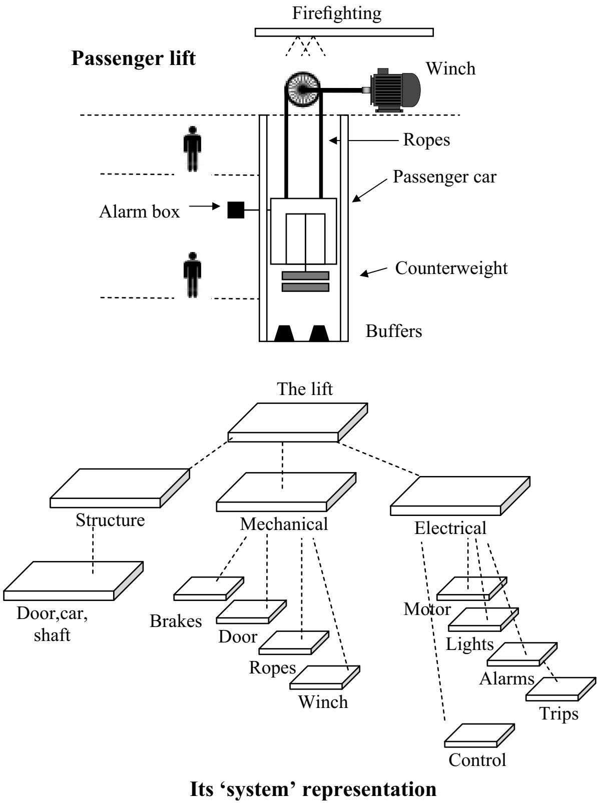


Figure 3.27 Representing a design as 'systems'. Reprinted from *Case Studies in Engineering Design*, Clifford Matthews, © Elsevier 1998

together to make a design 'work'. The methodology of this should be easier to understand by studying Fig. 3.27 – a simple passenger lift showing the design broken down using this type of systems approach. The lower part of the figure shows the function of the lift split into three primary systems: structure, mechanics and electrics. This primary system allocation is the most important, so note two points:

- Are you comfortable with the way that the lift shaft and car *structure* is shown as a system?
- A ‘system’ does not have to be a process, or an electrical or control network. Structures and mechanical components can also be thought of as systems.

Each primary system is then subdivided into its constituent subsystems, which gives a better resolution of what each subsystem actually does. In theory, you could go on indefinitely subdividing systems down the levels – a good practical approach. For the purposes of looking at embodiment design, it is best to use no more than three levels, as further subdivisions will overcomplicate the analysis.

Step 2: Consider redundancy
Try to think about redundancy before you get too involved with the details of individual parts of a design. The best technique is to list the various options, showing how they can apply to each system. Figure 3.28 shows a sample for the passenger lift design — note how it incorporates each of the four different types of redundancy shown previously in Fig. 3.25.

Step 3: List the danger features
You have to do this system by system. Any attempt to short-circuit the exercise by trying to ‘home in’ intuitively on the danger features that you feel are most obvious, or think of first, will not give the best results. Its steps need to be a systematic exercise — this is the whole key to opening up the problem and being able to identify *all* the design safety features, not just the easy ones. Figure 3.29 shows a typical analysis of danger features for one of the systems of the passenger lift. First you have to identify a danger feature, then do something about it. Now is the time to introduce the principles of design safety discussed earlier: avoidance features, protective features and warnings; in this order of preference.

Passenger Lift	A typical redundancy evaluation			
	Active	**Passive**	**Selective**	**Principle**
Shaft structure	High FOS	-	-	-
Car structure	**High FOS**	-	-	-
Winch	High FOS	-	-	-
Brakes	Multiple brake pads	Centrifugal brake	-	Electrical and centrifugal
Ropes	Multiple ropes	-	-	-
Doors	Parallel doors	-	Door closing	-
Motors	Design margins	Standby motor	-	-
Lights	Multiple lights	**Emergency lights**	-	Mains and battery
Alarms	Multiple sensors	Standby electrics	**Fire alarms**	Smoke and fire alarms
Trips	Multiple microswitches	Parallel circuits	Selective circuits	Mains and battery
Control	Multiple circuit	Parallel circuits	Selective positions	Electronic and hard-wired

FOS = Factor ofsafety

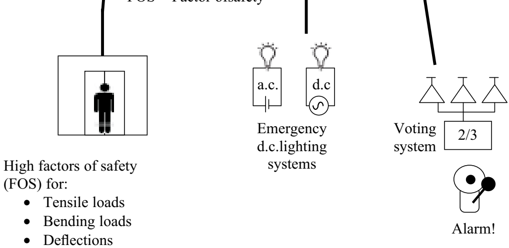


Figure 3.28 A redundancy evaluation. Reprinted from *Case Studies in Engineering Design*, Clifford Matthews, © Elsevier 1998

Good, objective embodiment design is about eliminating design safety problems at source, before they get *into* the design. It is a proven fact that once a feature becomes an accepted part of a design, and progress is made through the detailed engineering stage, it becomes very difficult to change it, without having to make changes to other (probably desirable) design features also. Natural reaction is to leave the original design features in, hence the only way to cover up the dangerous design feature is by *protection* — a less than ideal method, as we have seen.

Passenger lift door system:Embodiment/safety design features							
Design feature		***Avoidance***		***Protection***			***Warnings***
		Fail-safe	***Safe-Life***	***Control***	***FOS***	***Protection system***	
Stored energy	Door closing force	Soft door seals	Slow closing	Manual opening facility	Low closing force	Proximity interlock for doors	Yes
Rotating parts	External pulleys	Pulleys outside car	NA	NA	NA	Maintenance interlock	No
Crush and trap gaps	Between doors/top & bottom	<5mm gaps	NA	Manual opening	NA	Proximity interlock for doors	Yes
Exposed electrics	Behind car panel only	Low voltage only	Fuses	Controls all automatic	NA	Interlocked electrical doors	Yes
Hot parts		NA	NA	NA	NA	NA	No
Falling/slipping	Slips inside car only	No access to shaft	NA	NA	NA	Non-slip flooring and handrail	No
Noise	Low noise levels only	Remote winch	NA	NA	NA	NA	No

NA: Not applicable
Safe-life criterion comprises: Proven principles/calculations, assessment of operating limits and analysis of overload conditions

Figure 3.29 Embodiment design: danger features. Reprinted from *Case Studies in Engineering Design*, Clifford Matthews, © Elsevier 1998

The message is simple: danger features must be designed *out* at the embodiment stage, rather than trying to cover them up later. You can infer the basic methodology by a close look at Fig. 3.30. Although there are no rigid rules, you can generally get the best results by looking at the features in the order shown (i.e. working left to right across the table). Note that the figure only shows the analysis for one of the systems – for a full embodiment analysis, the process would be repeated for all the other systems identified previously.

Step 4: Look for embodiment options

It is rare that the first embodiment design ideas will be the best ones. The principle of finding the best solution lies with the activity of 'opening up' the design – revealing its complexity – to find the most appropriate solution from the 'possibles' available. Although the embodiment design activity benefits from a certain level of innovative thinking, it is important not to confuse this with true innovation. True innovation belongs in the conceptual stage that *precedes* the embodiment design steps – embodiment is a more prescriptive, better-defined process than this. Innovative embodiment solutions are fine, as long as they fit within the constraints of using those conceptual design decisions that have already been made. To reinforce this point, here are two examples relating to the passenger lift in Fig. 3.27:

- The conceptual design is for a 'rope operated' lift, not an alternative design operated by hydraulic rams. Hence the embodiment design should accept the use of an electric winch and look for ways to make the electrical system safe. It would be wrong to suggest the use of a hydraulic lifting system instead. That would be interfering with the agreed concept design.
- If, for example, the passenger lift was designed with double sets of doors, the embodiment design stage should accept this fact and look for options to make the double-door arrangement safer, not ways to change the concept to one using only a single set of doors.

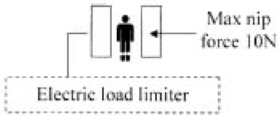
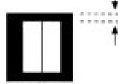
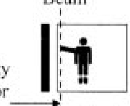
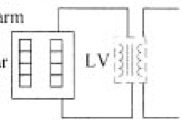
Danger	Embodiment	Good points	Bad points	Alternative
Stored energy. Door closing force		Simple electric limiter	Error signals (continual opening of doors)	Variable closing force
Crush and trap gaps		Safe: no further guards required	Can jam. Accurate assembly needed.	Cover-plates. Alternative guard arrangement
Crush and trap gaps		No moving parts	Unreliable sensors easily damaged	Pressure pads/ physical barrier
Exposed electrics		No HV danger.LV works OK	Transformer needed. More complex wiring	Use HV but include protection devices

Figure 3.30 Embodiment design options. Reprinted from *Case Studies in Engineering Design*, Clifford Matthews, © Elsevier 1998

This is a practical constraint to help to avoid the design process descending into a state of anarchy, not a mechanism to discourage innovation.

Figure 3.30 shows typical embodiment design options for the passenger lift, looking at the safety problem of crush/trap gaps between and around the lift's inner sliding doors. Note how sketches are used rather than detailed written descriptions for two reasons: (1) sketches help *definition*, they can capture and define that fleeting idea in a way that description cannot; and (2) sketches are a better way to *communicate* embodiment design ideas to other people. Other viewpoints are often necessary to help to 'firm up' ideas, but it is always better to use sketches.

- *Embodiment design is about specific features, not general principles – so use sketches.*

Step 5: Deciding

It is no use defining lots of elaborate design options if you can't then decide which to use in the final detailed design. A gamut of terminology surrounds this step; you will see it referred to as 'design evaluation' or 'design synthesis' through to the more elaborate term, 'evaluating concept variants'. They all mean the same: *deciding*. But how? Sadly, there are no hard and fast rules, it would be nice if there were. Several factors impinge upon the decisions, you can analyse and weight design options in order of safety, discuss them, and eliminate the worst ones – but you will still need to apply some intuition and experience. A few broad guidelines should be followed:

- ***Avoid contradictions* –** Make sure the technical choices are consistent with each other. There would be little logic, for example, in designing one electrical circuit for low-voltage operation, for safety, if others nearby operated on high voltage. Avoid contradictions by aiming for consistency.
- ***Use technical standards* –** These are useful to help you to decide. Design details in published technical standards have invariably been subjected to long discussions between people at the sharp end of manufacture and use of the equipment in

question. You can therefore rely on standards to provide *proven advice*. The main limitation is in the scope of technical standards – not all technical standards cover embodiment design, some being intended more as purchasing and specification guides rather than a design tool. This means that some technical standards are more useful than others.

- ***Technical guidelines*** **–** These are available in many forms: databooks, nomograms and manufacturers' publications. The quality of such information varies widely; databooks can be particularly useful for embodiment design ideas. Manufacturers' catalogues are good at showing different embodiment designs that are available (itself an indication of the success of the design feature) but tend to be optimistic in underestimating any negative aspects of a design option.
- ***Cost*** **–** More specifically, *cost effectiveness*. Frankly, you have to develop an instinct for cost as it is a real constraint in all design projects. Features such as redundancy and diversity are always desirable, but it is not economic to duplicate everything. Your objective should be to keep a focus on cost-effectiveness when deciding embodiment design – but keep it in perspective.

These aspects of deciding embodiment design should be treated as guidelines only. They must be seen as lying within the overall objective of making embodiment design as safe as possible, but also simple. Simplicity is a desirable design characteristic: good, safe, designs are often very simple.

3.12 Design by nature – project toucan

The relationship between nature and intelligent design has always been an awkward one. Many intelligent designs (engineered by humans to fulfil some technical necessity) follow those that can be seen to work well in nature. The spiral 'volute' shape of a pump casing is identical to that of the Nautilus seashell casing and the shape and structure of aircraft wings have a great similarity to those of birds. It's not all a procession after nature, however – nature has never invented the wheel and of the 118 elements in the periodic table, nature is only capable of making 93 of them occur naturally – the rest have to be manufactured artificially.

Overall, however, when searching for good design ideas, it is worth while looking at nature. Natural materials can produce interesting mechanical properties – spiders' webs, for example, have immense tensile strength per unit cross-sectional area, and the bite strength of a crocodile's jaw is in excess of that obtainable from a hydraulic ram, which requires huge tensile and compressive strengths from its various components.

Take the case of the toucan's beak (Fig. 3.31). This is one of the longest in the bird world, comprising more than 30% of the overall length of the bird. Impressively sized and coloured it performs remarkably in day-to-day toucan-like activities such as peeling fruit, fighting off predators and competitors, and impressing other toucans.

Natural design has been hard at work on this design, matching its mechanical properties to its needs, much the same as you would do for an engineering product made of metals. Consider these properties.

***Lightweight* –** Although making up 30% of the length of the body, a toucan's beak is responsible for less than 5% of its overall body weight. Weight is at a premium – birds require hollow bones and an ultra-lightweight skeletal structure if they are to be able to fly.

***Hardness* –** High hardness (resistance to surface indentation) is required for the usual toucan activities of pecking, fighting, peeling fruit, and splitting nuts. Hard surfaces can be made sharp, an advantage for all of these applications.

***Toughness* –** Toughness is the ability to resist brittle fracture under impact. As with metals, materials which are hard often have a tendency to be brittle. By their function, bird beaks are high-impact items, so need to be tough if they are not to break off in use.

***Tensile strength* –** Toucans use their beaks as levers to move stones, prise open gaps in bark and dig the occasional hole. Yield and tensile strength are important but so is stiffness – a high Young's modulus is required to keep deflections to a minimum.

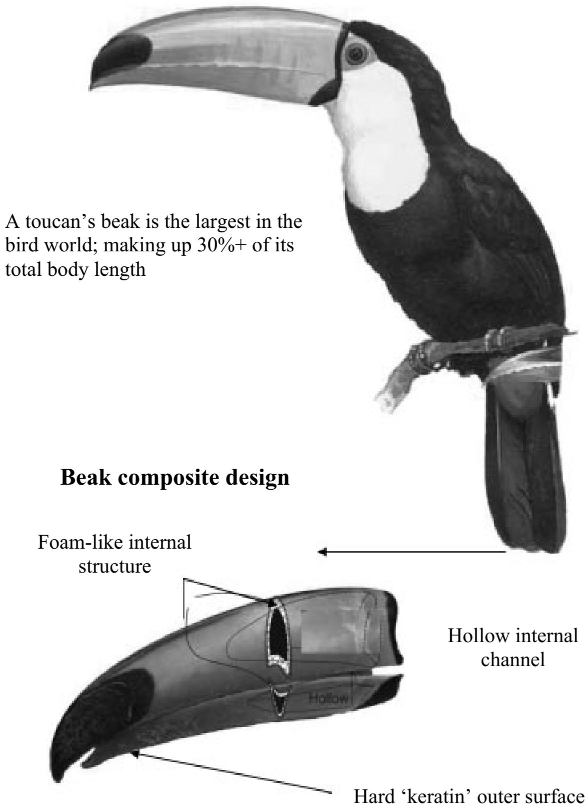


Figure 3.31 The impressive beak of the Toucan

Compressive strength **–** Rather than pure compressive 'crushing' strength as such, this is more to do with resistance to buckling. The beak must have a stable structure to stop it deforming and bucking (either axially or circumferentially).

Heat transfer **–** Flying is a high-energy business, generating a lot of heat that has to be transferred away to prevent overheating. A toucan's beak acts as one of nature's most efficient and adjustable radiators, capable of radiating almost 100% of generated heat when in flight. In many materials, heat transfer capability does not sit easily with strength – good conductors such as copper, aluminium and silver are quite weak materials even when alloyed with trace elements to improve their properties.

The solution: composite design

Almost the only solution to the toucan's list of design requirements lies with composite design. All natural designs are composites, of sorts, but this has to be one of the most striking examples. The internal structure of the beak consists of a closed-cell spaceframe containing pressurized air. This three-dimensional honeycomb construction is very stiff, providing resistance to buckling and torsion (twisting). It also provides resistance to impact – if a small crack were to start as a result of impact, it would soon be arrested by the convoluted path of the closed-cell honeycomb.

In more formal mechanics terms, the closed-cell structure provides an elastic foundation which increases its buckling load under flexure. It is made of fibres of a material called keratin, with mineral trace elements. These increase the strength, giving it a tensile strength of about 50–60 MPa and a high Young's modulus of 2.5–3 GPa.

For hardness, the beak is covered with overlapping keratin scales about 1 micron thick and 50 micron diameter glued together. As well as being hard, it achieves toughness, (resistance to brittle fracture) by allowing the glue to slip under impact. Together with the toughness of the underlying spaceframe, this allows the beak to absorb the worst toucan-induced impacts.

Finally, the hard shell maintains the underlying structure in slight compression, slightly deforming its multidirectional spaceframe struts, hence adding to its stiffness and resistance to torsion and buckling. In return, the foam-like spaceframe supports the shell, fortifying its lower stiffness (Young's modulus about 1–1.5 GPa).

This is a true example of composite design – two or more separate elements, each acting not only to their own strength but also *synergistically* to complement the other. There are many areas of design where composites can provide the combination of properties that a single material cannot. Classic examples are glass-reinforced plastic (GRP), carbon fibre and similar substances.

Thinking about it more widely, all useable metals are effectively composites – you just have to move down to the micro scale to see it. At this level trace elements, acting singly or together, enhance mechanical and chemical properties in fairly predictable ways. Much of this is hidden behind the scenes of course, and it is only relatively recently that the advantage of large 'touch and hold scale' composites have begun to be realized.

And behind it all?

Nature of course, has no intelligent designer – it is all done by natural selection sprinkled with a bit of random chance. If, perchance, there was an intelligent designer, then project toucan must have all gone wrong at the design review stage. Given that the easiest way round a design problem is to design out the need for it, the least effort solution would have been to simply give Mr Toucan a smaller 'off the shelf' beak. No problem with that, given of course, the control of the intelligent designer over the size and squishiness of Mr Toucan's favourite fruit, the aggressiveness of other toucans, and the inherent preferences of lady toucans. Eventually, with intelligent design in full methodical and analytical flow, all intelligently-designed birds would eventually look and perform almost precisely the same. Look at Formula 1 racing cars, and you get the idea.