Albert Cherns

# Principles of Socio-Technical Design[1][2]

The art of organization design is simultaneously esoteric and poorly developed. Most existing organizations were not born but "just growed." Many bear the recognizable stigmata of the operations of various well-known consultancy groups. There is, of course, no lack of available models, and no one seeking to set up an organization need reinvent the wheel. But organization design is generally an outcome, not an input. The input in manufacturing organizations is provided by the engineers, both those who design machines and equipment and those who design work methods and layout--the industrial engineers. Increasingly, operations researchers, systems analysts, the designers of computerized information systems, and the providers of "management services" of all kinds are having their say. In non-manufacturing work organizations, it is the latter who are most influential. And all of them, whether they recognize it or not, bring assumptions about people into their operations and their design. Most simply put, these assumptions can generally be described as Taylorist or System X--people are unpredictable; if they are not stopped by the system design, they will screw things up; it would be best to eliminate them completely but, since this is not possible, we must anticipate all the eventualities

---

[1]Slightly revised from a paper in Human Relations, 29:783-92, 1976.

[2]I am indebted to Louis E. Davis, on whose work in designing new organizations I have drawn heavily in this article, which arises out of the courses we have given together at UCLA and elsewhere.

and then program them into the machines. The outcome is the familiar pattern of hierarchies of supervision and control to make sure that people do what is required of them, and departments of specialists to inject the "expert" knowledge that may be required by the complexities of manufacturing, marketing and allied processes but is equally often required to make the elaborate control, measurement and information systems work.

We have found in our own work, in both teaching and consulting, that engineers readily perceive that they are involved in organization design and that what they are designing is a socio-technical system built around much knowledge and thought on the technical, and little on the social, side of the system. There is, of course, the danger that the term socio-technical system very rapidly becomes a shibboleth, the mere pronouncing of which distinguishes the cognoscenti from the ignorant and uninitiated. But recognizing that a production system requires a social system to integrate the activities of the people who operate, maintain and renew it; account for it and keep it fed with the resources it requires and dispose of the products does nothing by itself to improve the design. And while discussion of the characteristics of social systems is helpful, that still leaves us with the problem that there are many ways of achieving their essential objectives.

We teach engineers that any social system must, if it is to survive, perform the function of Parsons' (1951) four subsystems. As we present them, these functions are:

- attainment of the goals of the organization;
- adaptation to the environment;
- integration of the activities of the people in the organization, including the resolution of conflict, whether task-based, organization-based or interpersonally-based;

• providing for the continued occupation of the essential roles through recruitment and socialization.

The advantage of this analysis is that it tells designers that if they do not take these absolute requirements of a social system into account, they will find that they will be met in some way or other, quite probably in ways that will do as much to thwart as to facilitate the functions for which the designers plan. But it still leaves wide open the question of how to design a social system or, more fundamental, how much a social system should be designed. That there is a choice in such matters can be as much a revelation to the engineer as the fact that there is a choice of technology to achieve production objectives is to the social scientist.

How, then, do you design a socio-technical system? Can we communicate any principles of socio-technical design? The first thing to be said is that a lot depends upon your objectives. All organizations are socio-technical systems; that is no more than a definition, a tautology. But the phrase was first used with, and has acquired, the connotation that organizational objectives are best met not by the optimization of the technical system and the adaptation of a social system to it, but by the joint optimization (Emery, Vol. II, "The Nine-Step Model") of the technical and the social aspects, thus exploiting the adaptability and innovativeness of people in attaining goals instead of over-determining technically the manner in which these goals should be attained.

It is an obvious corollary that such design requires knowledge of the way machines and technical systems behave and of the way people and groups behave. Unless a designer is an engineer and a social scientist, both are required, which means engineers discussing alternative technical ways of attaining objectives with social scientists. This is not easy unless social scientists will take the trouble to learn enough about technology to understand the kinds of options that are open to engineers. The design team has indeed to be a multifunctional one as we have described elsewhere (Cherns, 1973).

In the process of designing ideas, no doubt the constant interchange among engineer, manager, social scientist, financial controller, personnel specialist and so on can do much to ensure that all aspects are considered, but the socio-technical concepts involved need not be hammered out afresh every time. They can be collected and presented in such a way as to ensure that they are taken account of without straitjacketing the designer. To this end, we have described nine principles, which we offer as a checklist, not a blueprint. They represent a distillation of experience and owe more to the writings of others (Emery, Vol. II, "The Nine-Step Model"; Emery and Trist, 1972; Herbst, 1974/Vol. II, "Designing With Minimal Critical Specifications") than to our own originality. They have not, however, previously been systematized.

## Principle 1: Compatibility

The process of design must be compatible with its objectives. A camel has been defined as a horse designed by a committee, and that joke unkindly incorporates negative evaluations of camels and committees. Camels certainly have minds of their own, but perhaps any attempt to draw more parallels between a camel and a social system would be unduly fanciful. Would a horse be more acceptable to a despot and a camel to a democrat? The point to be made, however, is that a participative social system cannot be created by fiat.

If the objective of design is a system capable of self-modification, of adapting to change and of making the most use of the creative capacities of the individual, then a constructively participative organization is needed. A necessary condition for this to occur is that people be given the opportunity to participate in the design of the jobs they are to perform. In a redesign of an existing organization, the people are already there. A design for a new organization has, however, to be undertaken before most of the people are hired. To some extent their jobs will have been designed for them in advance but this extent can be kept to a minimum. In one case (Davis and Cherns, 1975), the design team took the view that they would not design other people's lives. Having defined what were to be the objectives to be met and the competencies required to meet them, they deferred, until the individual was appointed, any discussion of how the job was to be performed. As in most cases "job" was not defined, this meant involving the people appointed as a team. Clearly some decisions had and have to be taken in advance; there has to be a pretty firm notion of how many people will be required and of what kinds of competence must be sought, but this is governed by the second principle.

## Principle 2: Minimal Critical Specification

This principle has two aspects, negative and positive. The negative simply states that no more should be specified than is absolutely essential; the positive requires that we identify what is essential. It is of wide application and implies the minimal critical specification of tasks, the minimal critical allocation of tasks to jobs or jobs to roles and the specification of objectives with minimal critical specification of methods of obtaining them. While it may be necessary to be quite precise about what has to be done, it is rarely necessary to be precise about how it is to be done. In most organizations there is far too much specificity about how and, indeed, about what. Any careful observer of people in their work situation will learn how people contrive to get the job done despite the rules. As the railwaymen in Britain have demonstrated, the whole system can be brought to a grinding halt by "working to rule." Many of the rules are there to provide protection when things go wrong for those who imposed them; strictly applied, they totally inhibit adaptation and even effective action.

In any case, it is a mistake to specify more than is needed because by doing so options are closed that could be kept open. This premature closing of options is a pervasive fault in design; it arises not only because of the desire to reduce uncertainty, but also because it helps designers to get their own way. We measure our success and effectiveness less by the quality of the ultimate design than by the quantity of our ideas and preferences that have been incorporated into it.

One way of dealing with the cavalier treatment of options is to challenge each design decision and demand that alternatives always be offered. This may result in claims that the design process is being expensively delayed. Design proposals may also be defended on the ground that any other choice will run up against some obstacle, such as a company practice, a trade union agreement or a manning problem. These obstacles can then be regarded and logged as constraints upon a better socio-technical solution. When they have all been logged, each can be examined to estimate the cost of removing it. The cost may sometimes be prohibitive but frequently turns out to be less formidable than supposed or than the engineer has presented it to be.

## Principle 3: The Socio-Technical Criterion

This principle states that variances, if they cannot be eliminated, must be controlled as near to their point of origin as possible. We need here to define variance, a word much used in socio-technical literature. Variance is any unprogrammed event; a key variance is one which critically affects outcome. This might be a deviation in quality of raw material, the failure to take action at a critical time, a machine failure and so on. Much of the elaboration of supervision, inspection and management is the effort to control variance, typically by action that does less to prevent variance than to try to correct its consequences. The most obvious example is the inspection function. Inspecting a product, the outcome of any activity, does not make right what is wrong. And if this inspection is carried out in a separate department some time after the event, the correction of the variance becomes a long loop, which is a poor design for learning. The socio-technical criterion requires that inspection be incorporated with production where possible, thus allowing people to inspect their own work and learn from their mistakes. This also reduces the number of communication links across departmental boundaries (see also Principle 5). The fewer the variances that are exported from the place where they arise, the fewer the levels of supervision and control that are required and the more "complete" the jobs of the people concerned to whom it now becomes possible to allocate an objective and the resources necessary to attain it. Frequently what is required to attain this objective turns out to be the supply of the appropriate information (see Principle 6).

Identifying variances and determining the key variances is a process often requiring lengthy analysis, and from time to time efforts have been made to codify it. One version, known as the nine-step analysis, has been developed by Davis and Cherns (1975). It has been used in enough organizations to give us some assurance that it can be adapted to use with any type of work organization, not just with manufacturing industry.

## Principle 4: The Multifunctionality Principle--Organism vs. Mechanism

The traditional form of organization relies very heavily on the redundancy of parts. It requires people to perform highly specialized, fractionated tasks. There is often a rapid turnover of such people, but they are comparatively easily replaced. Each is treated as a replaceable part. Simple mechanisms are constructed on the same principle. Disadvantages arise when a range of responses is required, that is, when a large repertoire of performances is required from the mechanism or the organization. This usually occurs if the environmental demands vary. It then becomes more adaptive and less wasteful for each element to possess more than one function. The same function can be performed in different ways by using different combinations of elements. There are several routes to the same goal--the principle sometimes described as equifinality. Complex organisms have all gone this route of development. The computer, for example, is a typical multifunctional mechanism. The principle of minimal critical specification permits the organization to adopt this principle also.

## Principle 5: Boundary Location

In any organization, departmental boundaries have to be drawn somewhere. Miller (1959/Vol. II, "Technology, Territory and Time: The Internal Differentiation of Complex Production Systems") has shown that such boundaries are usually drawn so as to group people and activities on the basis of one or more of three criteria: technology, territory and time.

Grouping by technology is typically seen in machine shops, where all the grinding machines are in one room, the Grinding Department; the milling machines in another, the Milling Department; and so on, with each department under the supervision of a specialist, a foreman grinder, etc. The consequences of this for the scheduling of work has been well described by Williamson (1972). A part in construction may be for months shuffled between departments, spending 1 percent of that time actually in contact with the machines. The consequent excessive cost of such work has been one of the stimuli to "group technology," the establishment of departments which each contain a variety of machines so that a part can be completed within one department. This corresponds to a grouping on the basis of time; the contiguity in time of operations indicates that they may well be organized together. Group technology also has consequences for the operation of the department as a team, with its members taking responsibility for the scheduling of operations and possibly the rotation of jobs.

Other examples of grouping on the basis of technology but not, of course, group technology, are the typing pool and the telephone switchboard. The switchboard may also be an example of the criterion of territory. Switchboard operators are bound together by the design of the machine. But the territorial principle can operate on the basis of little other than spatial contiguity. If the engineers have, for convenience, located different activities in the same area, the maintenance of control over the people working there suggests that they be made answerable to the same supervision. Retail trade organization is often of this kind with a floor supervisor. Organizations of this kind give rise to "dotted-line" relationships of functional responsibility.

All these criteria are pragmatic and defensible up to a point. But they possess notable disadvantages. They tend to erect boundaries which interfere with the desirable sharing of knowledge and experience. A simple example may suffice. In an organization concerned with the distribution of petroleum products studied by Cherns and Taylor (unpublished), the clerks who collected customers' orders were organized in a department separate from that of the drivers for whom schedules were worked out. A driver would pick up a schedule allocating him a vehicle and a route. Frequently the receipt of the routing would stimulate a string of expletives from the driver: "If I did what this *** has told me to do, I should not be able to do half the job. I would arrive at customer B just after 12 o'clock when the only man with the key to the pumps has gone off on his lunch break. And it's no use my turning up to customer P until I have discharged enough of my load for his short pipe to reach my tank. And finally I would end up on the *** road just in the middle of the rush hour. It would serve him right if I followed these instructions. I would run out of time [exceed the permitted number of consecutive driving hours] right in the middle of a throughway." There was no doubt pardonable exaggeration in all this; the point is that the drivers had acquired a great deal of knowledge about customers, routes, etc., but being organized into a separate department, they shared very little of this knowledge with the routing clerks who, however, received the customers' complaints before the drivers.

The principle has certain corollaries. A very important one concerns the management of the boundaries between department and department, between department and the organization as whole and between the organization and the outside world. The more the control of activities within the department becomes the responsibility of the members, the more the role of the supervisor/foreman/manager is concentrated on the boundary activities--ensuring that the team has adequate resources to carry out its functions, coordinating activities with those of other departments and foreseeing the changes likely to impinge upon them. This boundary maintenance role is precisely the requirement of the supervisor in a well-designed system.

Under favorable circumstances, working groups can acquire and handle a greater degree of autonomy and learn to manage their own boundaries. This implies locating responsibility for coordination clearly and firmly with those whose efforts require coordination if the common objectives are to be achieved. The role of supervisor now becomes that of a "resource" to the working group.

## Principle 6: Information Flow

This principle states that information systems should be designed to provide information <u>in the first place</u> to the point where action on the basis of it will be needed. Information systems are not typically so designed. The capacity of computer-controlled systems to provide information about the state of the system, both totally and in great detail, to any organizational point has been used to supply to the top echelons of the organization information that is really useful only at lower levels and that acts as an incitement to the top management to intervene in the conduct of operations for which their subordinates are and should be responsible. The designer of the information system is naturally concerned to demonstrate its potentialities and is hard to convince that certain kinds of information can be potentially harmful when presented to high organizational levels. Properly directed, sophisticated information systems can, however, supply a work team with exactly the right type and amount of feedback to enable them to learn to control the variances which occur within the scope of their spheres of responsibility and competence and to anticipate events which are likely to have a bearing on their performance.

## Principle 7: Support Congruence

This principle states that the systems of social support should be designed so as to reinforce the behaviors that the organization structure is designed to elicit. If, for example, the organization is designed on the basis of group or team operation with team responsibility, a payment system incorporating individual members would be incongruent with these objectives. Not only payment systems, but systems of selection, training, conflict resolution, work measurement, performance assessment, timekeeping, leave allocation, promotion and separation can all reinforce or contradict the behaviors which are desired. This is to say that the management philosophy should be consistent and that management's actions should be consistent with its expressed philosophy. Not infrequently, a management committed to philosophies of participation simultaneously adopts systems of work measurement, for example, which are in gross contradiction. Even management as progressive and committed to the humanization of work as that of Volvo's Kalmar plant has retained a commitment to a system of payment based on MTM, a technique of work measurement utilizing time and method study. Until replaced, this may pose an obstacle to the further humanization of work at Kalmar to which the management is committed.

## Principle 8: Design and Human Values

This principle states that an objective of organizational design should be to provide a high quality of work. We recognize that quality is a subjective phenomenon and that everyone wants to have responsibility, variety, involvement, growth, etc. The objective is to provide these for those who do want them without subjecting those who do not to the tyranny of peer control. In this regard, we are obliged to recognize that all desirable objectives may not be achievable simultaneously.

What constitutes human work is a matter again of subjective judgment based on certain psychological assumptions. Thorsrud (1972) has identified six characteristics of a good job which can be striven for in the design of organizations and jobs:

- The need for the content of a job to be reasonably demanding of the worker in terms other than sheer endurance, and yet to provide a minimum of variety (not necessarily novelty).
- The need to be able to learn on the job and to go on learning (again, it is a question of neither too much nor too little).
- The need for some minimal area of decision making that individuals can call their own.
- The need for some minimal degree of social support and recognition in the workplace.
- The need for individuals to be able to relate what they do and what they produce to their social life.
- The need to feel that the job leads to some sort of desirable future (not necessarily promotion).

## Principle 9: Incompletion

Design is a reiterative process. The closure of options opens new ones. At the end we are back at the beginning. The Forth Bridge, in its day an outstanding example of iron technology, required painting to fend off the rust. Starting at the Midlothian end, a posse of painters no sooner reached the Fife end than the Midlothian end required painting again. Varying the image, Jewish tradition prescribes that one brick be omitted in the construction of a dwelling lest the jealousy of God's angels be excited. Disregarding the superstition, the message is acceptable. As soon as design is implemented, its consequences indicate the need for redesign. The multifunctional, multilevel, multidisciplinary team required for design is needed for its evaluation and review.

## Concluding Remarks

Who is the socio-technical designer to whom this paper is especially addressed? The analysis, preparation and implementation of a socio-technical design is the property of no individual or set of individuals; it belongs to the members of the organization whose working lives are being designed. Special skills and knowledge may well be, and often are, required, and these are provided as a resource by socio-technical consultants or action researchers.

But participation by employees in the design of their organizations may imply that they accept, or show readiness to accept, work roles which go beyond the agreements and constraints evolved by negotiation between management and union on their behalf. Unions are thus inevitably involved in the process, whether in a collaborative, neutral or antagonistic role. Can they be partners in design? This is a role which has seldom been offered to, and even more rarely accepted by, unions. It is not a role for which they have prepared themselves, and it is one which could easily blur their primary responsibilities to their members. Yet without them the viability of the design must be in some doubt. And the design of a social support system implies designing the functions of the shop steward, if not the union official. Our first principle, compatibility, requires that the unions be brought into the design if that is at all possible. But if they are to come in, they, too, will need to acquire new competencies.

## References

Cherns, A.B. 1973. "Helping Managers: What the Social Scientist Needs to Know." Organizational Dynamics, Autumn:51-67.

Davis, L.E. and A.B. Cherns. 1975. The Quality of Working LIfe, Vols. 1 and 2. New York: Free Press.

Emery, F.E. and E.L. Trist. 1972/73. Towards a Social Ecology.London/New York: Plenum Press.

Herbst, P.G. 1974. Socio-Technical Design: Strategies in Multi-Disciplinary Design. London: Tavistock Publications.

Miller, E.J. 1959. "Technology, Territory and Time: The Internal Differentiation of Complex Production Systems." Human Relations, 12:243-72 Voll II, pp. 385-404.

Parsons, T. 1951. The Social System. London: Routledge and Kegan Paul.

Thorsrud, E. 1972. "Policy Making as a Learning Process." In Social Science and Government: Policies and Problems, edited by A.B. Cherns, R. Sinclair and W.I. Jenkins. London: Tavistock Publications.

Williamson, D.T.N. 1972. "The Anachronistic Factory." Proceedings of the Royal Society, A331:139-60.