13 How Do We Know It Works? Approaches to the Evaluation of Complementary Medicine

Some starting points

Complementary medicine is a term used to cover a vast array of treatment procedures as wide ranging as aroma therapy, iridology, acupuncture, homeopathy and osteopathy. It is sometimes known as alternative or even folk medicine. These complementary therapies exist because people find them helpful. Research must, therefore, be directed at asking 'Who?', 'What?', and 'How?'. In the following chapter ways of answering these questions are examined. Emphasis is given to the need for a variety of different research strategies and tactics. A preliminary outline of an overall approach which would hold together the strands in a fruitful way is proposed, based on facet Meta-Theory.

In considering the emerging research prospect for the study of complementary medical practice it is important to clarify some initial premises.These provide a framework within which to consider a great variety of research possibilities.They also point towards areas of study that might not otherwise be apparent.

Complementary medical practice exists and thrives

Fulder's (1984) review of over 250 organisations and colleges for practitioners, taken together with the more recent survey by Wharton and Lewith (1986), showing the extensive use made of those services through referrals from medical general practitioners, serves to illustrate the vast need there seems to be for complementary medicine.

This fast growing network of alternative services means that research cannot proceed on the basis that there are some novel untried procedures that must be proven, before they can be released to an unsuspecting public. The researcher is not faced with the recent discovery of some purported new cure that will not be taken up unless research proves that it works. The situation is more like that of a committed, scientific atheist, who wishes to understand the values of religion.

This chapter is derived from papers that were greatly influenced or indeed co-authored by Ceri Roderick and Lorraine Nanke. I am happy to acknowledge their contributions to the thinking behind the present chapter as well as their detailed research input.

Religions of many forms exist on a large scale. Their destructive potential cannot be denied, but many will speak up for their fundamental, essential value in society. The task of research is to understand the strengths and weaknesses of these existing approaches to the promotion of human wellbeing. To develop that understanding we must examine the logic upon which the different complementary practices operate and study closely the conditions under which successes and failures occur. Above all, the very existence of complementary medicine demands that research establishes exactly how and why it exists. Merely proving or disproving the specific effects of the agents (needles, pills or manipulations) used in complementary medicine, will have as little effect as Marx's arguments about the illogicality and evil of religion.

Metaphors for disease processes

The need to broaden our approaches to the study of complementary treatments is illuminated when it is realised that we cannot see diseases; they are an invented notion. We may see their consequences but not the disease process. So, in order to talk about a disease we have to develop ways of referring to those processes, particularly if we want to talk about possible outcome and treatment and what is being done to control or eradicate the disease. To do this, metaphors have to be used to describe what is going on.

For example, a common metaphor is warfare. We are 'attacking' the disease. This is one way of talking about the process between doctor and patient. In fact it is quite an interesting and rich metaphor. There are even local 'guerrilla battles' that are being fought. These contrast with, say, somebody in intensive radiotherapy or chemotherapy where it becomes a total war in which the individual accepts total involvement.

Another metaphor may be thought of as horticultural. The idea that there are natural processes that with enough light and effective watering and so on, will come to fruit. This is just as much a metaphor and just as necessary to use as any of the others.

Other metaphors that are very common are a variety of mechanical metaphors, not just in relation to the body. The idea that you can localise the seat of action of a particular antibiotic, for example, and thereby nudge the offending component out of the way in order to change the system. Another sort of metaphor that is related but rather different is the hydraulic metaphor of pushing forces along a pipeline. Acupuncture speaks of energy and forces and so uses these ideas overtly, but there are many other practices which are similar in their vocabulary. One that is often used in surgery, is the plumbing metaphor. It is often very powerful for the individual to know that his 'pipes are blocked' and that they can be unblocked.

There is another fascinating metaphor that is sometimes used today, yet it is difficult to find a word for it. Perhaps 'acrobatics' is a good term because it helps to explain why people are so hesitant in certain areas, especially in psychotherapy. The business of 'getting hold of yourself', of 'coming to terms with yourself', of

developing a way of thinking about yourself, about gaining control over your own actions, the whole issue of agency and the individual being a dominant performer in his own display, is often presented in a way that involves all sorts of psychic contortions, and on occasion also physical ones. That is one of the reasons why I think people find the psychotherapeutic encounter so difficult to understand, so threatening, because it does involve going through mental gymnastics that they might feel they are not up to.

One further metaphor is really close to magic and witchcraft, especially the idea of exorcism. This seems to be still a very powerful idea in many areas; the idea of driving the evil out, 'If you take these magic pills they will make the disease go away'.

I have deliberately chosen metaphors that are equally applicable to complementary and allopathic/conventional medicine. All are used as part of the process of negotiating and discussing the nature of the problem that the individual is facing.

Returning to the general framework, the problem of generating ways of measuring outcome and deciding whether treatment actually works, is a problem of finding measurement devices that will connect with the sort of metaphorical explanations that are at the heart of our dealing with patients.

The question 'How do we know it works?' is a very powerful mechanical metaphor. It enshrines the idea of a machine rolling on, and of bits falling off or being broken somewhere. We replace a part or fix it in some way. Then we ask 'How do we actually know whether it is still working?' We accept that the person is such a complicated machine that you cannot actually look at it and see that it is working in the same way that it was, and that in a few years time the weakness may show up again somewhere.

If we had asked the question 'How do we know that it has gone?', it would have been more of an exorcist form; 'How do we know that the individual has got it under control?' would be more acrobatic; 'How do we know when we are winning?' would be military; and 'How do we know that channels are unblocked, or fluids are in balance' a hydraulic form of question. Each has equal legitimacy. But each draws attention to different ways of considering the whole system of experiences that go to make up the treatment process. If research ignores this complexity and seeks only to examine bio-medical processes it may make some small-scale discoveries, but it will miss the overall picture.

Conventional research activities are frequently flawed

The first volume of the *Journal of Complementary Medical Research* was devoted to the strengths and weaknesses of the controlled, randomised experimental trial. It is clear from that discussion and many other considerations of research procedures that it would be unduly limiting to restrict complementary medical research to the experimental research strategy. But it is also important to realise that, as with all activities conducted by human beings, all research is subject to many different biases and flaws. Furthermore, any given study involving people as

the subjects of the study is more open to error than might often be realised. This is because the study of biological, psychological and social processes is particularly difficult and, especially where practitioners are carrying out the studies, because practitioners, even though sophisticated in theory, are often not expert in research methodologies. Many purely technical infelicities have been pointed out in the published literature on research on people, even before the philosophical and ethical problems involved are considered.

No one study or research strategy will reveal the untrammelled truth, and inevitably a variety of strategies will be needed. Attempts should be made to identify the biases in any one strategy and deliberately complement that approach with others that are not so open to that bias.

The scientist-practitioner model

One common approach to therapeutic research has been termed the scientist-practitioner: the clinician who not only helps relieve patients' problems using the expertise of their own discipline, but also carries out research into their own practice. The model is based on the idea of a continuing dialogue between theory and practice. Clinical work provides a source of insights and hypotheses about the effects and effectiveness of particular interventions. These are then subjected to more systematic evaluation in individual cases and larger scale group studies, eventually contributing to the development of collective knowledge and improvements in practice.

In relation to psychotherapy, Strupp(1981) suggests that the scientist-practitioner model provides: 'the opportunity for oscillation between the observation and participation, between taking part and standing back, between feeling and thinking, between (controlled) abandonment and study. It is this process of oscillation, the unique human ability to resonate, identify and therapeutically respond to themes in the patients' experience that is essential ... It should be possible to encourage research that is both rigorous and relevant to clinical and social issues' (p. 563).

Advantages of the model

Overcoming the clinical/research split. A major advantage of the model is that it overcomes the sterile split between clinicians and research scientists, potentially avoiding the narrowness of thought and rigidity of action which are often associated with exclusive focus on either. Scientific researchers are often criticised for producing results which seem irrelevant to clinical practice using measures and strategies which are beyond the scope and resources of clinicians. Conversely, clinicians are criticised for not using the insights acquired as a result of their experience as a basis for research, development of cumulative knowledge, or training, trusting solely on their clinical experience to evaluate their practice. The model is intended to overcome this problem by combining clinical and research functions in the same individuals where possible.

Accountability. The increasing stress on accountability in the provision of services means that particular therapeutic disciplines cannot avoid the necessity of describing and empirically validating the effects of their work. This requirement is likely to become increasingly important to patients who wish to make an informed choice between the wide diversity of treatments available. In this context, all clinical disciplines have an obligation to carry out some type of empirical research into their methods of treatment.

Lack of dependable knowledge. Perhaps the most important consideration underlying the scientist-practitioner model is the lack of dependable information currently available about the reliability and effectiveness of many diagnostic and therapeutic techniques. In 1950, Raimy stated that: 'psychotherapy is an undefined technique, applied to unspecified problems, with an unpredictable outcome. For this technique we are recommending rigorous training' (p. 93). It is disappointing to recognise how much this comment is still applicable to psychotherapeutic practice and it may be useful to consider the extent to which it is also relevant to complementary medicine. Clinicians of all disciplines are best placed to recognise and study the important aspects of real life clinical intervention and change; indeed most rely on their own experience as validation of their particular theoretical approach and techniques. Though such personal conviction may be valuable to the individual clinician and their patients, it is of little help in contributing to collective, dependable knowledge about the process and effects of different types of therapy. This requires systematic empirical investigation producing objective results which can be evaluated by others within the discipline and the wider scientific community.

Problems of the model

A salutary lesson from clinical psychology. Unfortunately, the scientist-practitioner model remains more of an ideal than an accurate description of normal practice, even in clinical psychology which is explicitly based on this approach. Despite a thorough grounding in research methodology, very few practising clinicians actually carry out research after qualification, and many state that they would prefer less emphasis on research methodology in training. Some consideration of this issue may help complementary practitioners avoid a key pitfall of the scientist-practitioner model, and help establish the research enterprise on a more productive footing.

The clinical context and experimental science

One of the major problems facing clinicians who wish to carry out research has been identified as the mismatch between the demands of clinical research, and the requirements of experimental science. This was the conclusion drawn by Bergin and Strupp (1981) from a survey of leading psychologists and psychiatrists. They state that: 'among researchers as well as statisticians there is a growing disaffection

from traditional experimental designs and statistical procedures, which are held inappropriate to the subject matter under study. This judgement applies with particular force to the area of therapeutic change, and our emphasis on the value of experimental case studies underscores this point. We strongly agree that most of the standard experimental designs and statistical procedures have exerted, and are continuing to exert, a constricting influence on fruitful inquiry, and they serve to perpetuate an unwarranted emphasis on methodology' (p. 440). This mismatch has often been dealt with by attempting to adapt clinical questions to the Procrustean bed of experimental science. As such, it is often accepted as axiomatic in medical, psychological and biological research that the only genuinely scientific strategy is the experimental design, in which people are randomly assigned to different conditions and subjected to blind or double blind trial in order to measure the influence of a key independent variable. The independent variable in such studies is characteristically seen as the main single causal agent producing the effect, which is the observed difference between groups. The limitations of this approach have been well documented elsewhere, (e.g., Heron, 1986) but it is worth noting some of the main difficulties encountered in a clinical context here.

Conflicting priorities

A clinician's first priority is almost always the wellbeing of individual patients, and this may conflict with the requirements of strictly experimental research design. For example, treatment strategy may normally be decided on the basis of ongoing therapeutic dialogue with the patient, changing in response to the patient's reaction, and reformulations of the problem. Yet, controlled comparison between two specific techniques often requires that the treatment, or lack of treatment in the case of a control group, is determined by the requirements of the research design rather than patient need. The scientist-practitioner model is likely to be most effective when the primary clinical priority is recognised, and means found to help clinicians carry out research into what they actually do in the course of normal practice rather than tailoring practice to the experimental design.

Moral concerns

Closely related to the issue of clinical priorities are the moral problems which can be encountered using the standard clinical trial. Firstly, there is the difficulty of obtaining patients' informed consent, as this kind of study is ideally carried out double blind, without patient or practitioner knowing which patient is receiving which treatment. Secondly, many clinicians express concern about patients in the control group being untreated, given a placebo, or receiving a treatment which may be less effective than the experimental treatment.

Single versus multiple factors

The experimental method is designed to address questions which can be answered

in terms of simple cause-effect mechanisms.The clinical trial is only able to compare simple alternative causal agents, i.e., the variable which differs between control and experimental groups, to explain observed variation in effect. This kind of question is perfectly appropriate where there are clear theoretical reasons to predict that treatment agent A will cause outcome B; for example, that a particular antibiotic will remove infection and associated symptoms, or that any difference in outcome is likely to be due to the single factor on which the control and experimental groups differ, e.g., type of treatment received. The clinical trial is particularly well suited to the traditional disease focus of orthodox medicine, in which the search for single pathogenic agents and remedies has a strong theoretical basis. However, it is less applicable to the holistic approach, which is based on a theoretical framework stressing the importance of multiple interacting factors relevant to treatment effects in each individual patient.

Clinical and statistical significance

Associated with the clinical trial methodology has been an over-reliance on statistical probability measures to compare group differences. These measures are often used inappropriately, and lead to impoverished interpretation of results, particularly in relation to small sample sizes. Statistical significance is frequently confused with clinical significance, and all too often studies report only data on the probability that differential treatment outcomes are a result of chance, and assume that the treatment which produces a reliably different outcome is better. Yet, it matters little if one therapy produces effects at a statistically significant level more than other therapies, if these changes are not of sufficient magnitude to be experienced as meaningful by the patient. Even statistically significant results often account for only a small proportion of the observed variance, indicating that in such cases the single cause explanation provides an inadequate account of obtained results.

Practical difficulties

The practical requirements of setting up clinical trials, including obtaining large numbers of matched subjects, control groups, possibly external blind raters, and large scale standardisation of treatment, can place excessive and possibly unmanageable demands on practising clinicians who might otherwise be prepared to contribute to research. Further, 'blind' administration can easily be applied to pharmacological treatments in which active and placebo substances can be made to appear identical. However, it is impossible to administer many other forms of physical or psychological treatment 'blind', for example massage or acupuncture; such treatments an be incorporated into standard clinical trials only if an apparently equivalent but theoretically inert form can be devised and administered to the control group.

Group versus individual differences

The practice of using group averages to compare patient groups obscures potentially important individual differences in treatment response. Kiesler (1965) has described this practice as based on 'the patient uniformity myth', that is, the assumption that the variables being studied are the only relevant differences between groups. This assumption is not warranted on empirical grounds, as there is much evidence that some patients in every study improve, some stay the same, and some get worse (see for example, Garfield, 1981). Barlow et al (1984), point out that the roots of the problem lie in traditional experimental research methodology, and particularly the improper use of sampling techniques.

According to sampling theory, a random sample of the population of interest must be studied if adequate generalisations are to be made about the whole population. In a clinical context, this would mean drawing a sample in which all individual differences potentially relevant to treatment outcome are represented. If the treatment applied to all of these individuals produced a statistically significant effect compared to a control group, the treatment would be judged effective. However, Barlow et al add that applied researchers have long recognised firstly the practical impossibility of obtaining a patient sample in which all potentially relevant factors are adequately represented, and secondly the theoretical problem that the sample would be so heterogenous that few if any treatments would be likely to show statistically significant effect.

An alternative strategy consistent with experimental methodology and the clinical trial has been the selection of homogenous patient groups to minimise sources of individual variation which may confound results. This approach is also subject to the practical problem of obtaining an appropriate sample; for example, finding a group of anxiety patients who also have similar personalities, backgrounds, life circumstances and any other factors of potential relevance to treatment. Further, the results of such studies can only be generalised to patients with similar characteristics, and so are of limited value to practising clinicians. What is needed is an approach which can go beyond the comparison of group averages, to help identify which patients respond in which way to which treatment.

Experimenter distortion

Experimenter distortion in research is a double-edged sword. On one hand, the most productive research is usually carried out by those with a strong interest and expertise in the subject area. On the other hand, science is based on the ideal of objectivity and freedom from personal bias. Though a study may be designed on the basis of personal interests, the results should be reproducible by other independent researchers with different interests. One potent way in which the researcher can distort or influence data obtained is by inadvertently conveying their expectations to patients. Orne (1962) suggested that subjects are usually so keen to please the researcher that they actively search for and act on clues as to how the researcher wants them to behave. These clues and hints are collectively referred to

as the 'demand characteristics' of a situation. 'Demand characteristics' are clearly involved in any clinical or research enterprise involving human beings. However, tailoring treatment to fit strictly experimental research design may so alter the perceived demand characteristics for patients that results obtained may not easily generalise to the different perceived demand characteristics of the normal clinical consultation. Experimenter distortion may be one reason for the fact that one of the few reliable generalisations that can be drawn from the wealth of psychotherapy outcome research is that most studies provide some degree of support for the orientation of the researcher.

Quantitative and qualitative data

One fallacy often associated with reliance on experimental methodology is the assumption that easily quantified factors are inherently more scientific than those which are not so easily assigned to numbers. This has been likened to the fable of the drunk looking for his key under the light rather than where he lost it, because he can see under the light. The important point here is a that a research study should be designed to assess the factors which are likely to be relevant to the original question, rather than those which are easily measured. This applies in particular to clinical research, where qualitative changes, for example sense of wellbeing or relationships, may be of overriding importance. The goal of objectivity and replicability is central to any type of measurement, but different types of measurement are appropriate for different types of variable. The task is to identify ways in which relevant factors can be most reliably assessed, rather than trying to build research questions around factors which can be easily quantified.

Temporal changes

Finally, many experimental studies compare treatment outcome at one particular point in time, classifying patients in terms of extent of recovery. The usefulness of this approach may depend partly on the condition being studied; for example it makes considerable sense in relation to acute conditions such as an infection, or emotional crisis reaction, but is perhaps less applicable to chronic or lifestyle-related conditions, in which longer term changes, and maintenance after treatment are more significant.

The need for a meta-theory

Once these assumptions are recognised, it is apparent that a plethora of unconnected non-cumulative studies is likely. In the complementary arena such diffusion of effort is even more likely because of the many different theoretical formulations from which hypotheses can be derived. The significant concern that variations between individuals should be fully accounted for serves only to add complexity to an already potentially confusing situation.

There are a number of ways out of this possible morass. One is to identify and focus attention on high priority areas, especially those with political significance, and focus attention on those. Unfortunately, priorities today may be irrelevant tomorrow. This is unfortunate in research terms because research initiated today will certainly not report until tomorrow, if not the day after. Another approach is to provide a framework within which the theories and methods of all research can be accommodated. Such a framework must be value-free and neutral as regards particular theoretical stances, but will allow comparisons between different studies.

This might seem a great deal to demand of any framework, until it is realised that what such a framework is being called on to provide is simply a coherent logical system for describing research. Such a system needs to provide the intellectual tools for summarising the particular theories under study and for establishing the links that those theories have with recorded data. In other words, the approach that will provide a framework for a variety of different studies, drawing on different theories, without forcing them all into the same theoretical straitjacket, will be a theory about how theories are constructed and tested, a meta-theory.

An effective meta-theory would enable researchers to cast, say, acupuncture and homeopathy into the same logical mould, together with conventional medicine without the need to interpret all the data as energy lines, or as pharmacological changes, or whatever. Of course, eventually it would be expected that one explanatory system would become dominant, because of its power and fruitfulness, but in the best scientific traditions there would be no need to assume, in setting up the research, which theory is the most powerful.

The level of abstraction of this argument is so great that it is worth giving a concrete example of the essential logic of this approach. Consider homeopathic prescribing on the basis of the identification of the patient as relating to a particular polycrest. A strongly theory-driven research project would, typically, try to find some existing classification of the patients being treated, say, in conventional medical terms and either try to 'validate' the homoeopathic prescription by relating it to known pharmacological effects, or by trying to convert the polycrest into the conventional medical diagnosis and extrapolate the results.

If a meta-theory can be drawn upon, however, the homoeopathic procedure can be examined initially in its own terms. It will reveal that certain relationships between what happens to patients, how they are initially classified and the treatment prescribed will be predicted by the theory. A conventional medical approach will also generate an interrelated set of predictions. Both approaches can therefore be examined to see which is the most fruitful in the consideration of the treatment and its outcome.

It is important to note that in the example no reference is made as to which treatment is the most 'successful'. Success is a higher-order level of analysis and may be made irrelevant by the earlier level of discourse, for example by the fact that the homoeopath, or GP, cannot specify what outcome should be characteristic of the treatment given. As this example shows, the meta-theory requires to be quite robust and logically powerful if it is to make any significant contribution.

A brief outline of facet meta-theory

The value of a meta-theory also depends upon how flexible and fruitful it proves itself to be. Mathematical procedures are generally the basis of any meta-theoretical approach, but in the area of complementary medicine many of the issues to be studied will not lend themselves readily to advanced mathematics. One approach that is gaining in acceptability, facet meta-theory, derives from elementary mathematical ideas. This has been applied to studies ranging from distinguishing malicious from non-malicious false alarms to the fire brigade to the prognosis of thyroid cancer, from research on the provision of hostels for alcoholics to explanations of slimming behaviour. In essence the facet approach sees any scientific theory as consisting of four components:

1. The first, and most central, is the recognition that all theories are in essence derived from ways of classifying observations, incorporating some clear definitions of what the classifications are. This is known in facet theory as a 'definitional system'.
   Darwin's theory is based on the classification of plants and animals, etc. Bacteriological theory is derived from categories of disease, symptoms and bacteria. Acupuncture is derived from classes of energy imbalance, osteopathy from classes of position of bones and so on. Any mutually exclusive, exhaustive set of categories is known as a 'facet'. Facets are assumed to be only qualities, unless there are particular reasons why some order or quantification can be given to the 'elements' of which they are composed. The colours of the spectrum make up the qualitative elements of a colour facet, unless they are defined in the quantities of wavelength. This fundamentally qualitative basis to facet theory is what gives it its power and flexibility.
2. The second component is the identification of patterns in any observations made. An elementary pattern would be the observation that those assigned to one treatment group all have measured levels that are distinct from those assigned to another group. The notion of a 'pattern' in the observations is, however, a very rich one that can accommodate notions as complex as multivariate relationships or as simple as differences in percentages.
3. The third component is the establishing of a correspondence between some aspect of the definitional system and some critical aspect of the patterns in the observations. People described as having been treated with a particular drug and whose responses have been classified according to how much better they are, would be expected to have a different sub-set of responses to those who were not described as having been so treated.
4. The fourth component is a logical explanation or 'rationale' as to why the correspondence between definitional system and observations occurs.

Definitional system, pattern in the observations, correspondence and rationale are all directly recognisable in any good science, whether it uses the language of facet theory or not. In that respect facet theory is simply a systematisation of the general

scientific approach. But by being so systematic and theory-tolerant, it does provide opportunities for research in complementary medicine that are hidden when less overt systems of science are utilised.

A general data matrix

In order to see how the facet approach can define a research agenda there is one helpful technical device, the data matrix.

In essence, all therapeutic research can be viewed as the exploration of patients and their attributes. The simplest way of thinking of this is that data is collected that can be represented as a number of rows, one row for each patient, and a number of columns, one column for each facet that characterises the patients. The cells of the matrix indicate the value any given patient has achieved on any given facet. For example, in its most elementary form a drugs trial of 20 people could produce a matrix of twenty rows and two columns, one classifying the patients in terms of the drug they received, the other in terms of its effects. The pattern looked for is one that matches the two classifications. If there is no discernible match then nothing clear can be said about the effect of the drug.

Matrixes of greater complexity can, of course, be constructed. They can mix together, for example, the many psychological and physiological variables that a holistic approach would suggest should be combined. The data matrix knows and cares nothing of such miscegenation! The rows need not be different people but can be the same person at different points in time, or some combination of the two. In other words, like the facets to which it relates, the data matrix makes no assumptions about the most appropriate theories or methods for the research in question. It is a neutral vehicle for carrying out research. But, as will be seen, it helps to define a number of otherwise rather vague research objectives. Of course, as the matrix becomes more complex so more demands are made on the analysis procedures to find the critical 'patterns'. New developments in computing and statistics greatly facilitate the matrix analysis, but elaboration on these more technical aspects is not appropriate here.

Objectives for research

Having established a context for considering complementary medical research and a framework within which to specify such research, it is now possible to elaborate some of the main objectives for research and indicate how that research might be structured.

Research priorities in complementary medicine

On this account of science, the most important task facing complementary medicine

is to define key research questions. This involves clarification of conceptual systems used in different disciplines into a form which can be evaluated by empirical data.

Reliability of clinical judgements

When measuring stable characteristics such as personality, measurements must be stable, consistent and free from random error or chance fluctuations. For example, if a different result was obtained each time a particular individual's height was measured, something would be seriously wrong with the measuring device. The issue is more complex with regard to unstable changing characteristics such as health status which are more often the subject of clinical concern. In such cases, reliability can best be assessed in terms of inter-rater agreement in clinical judgement, or the extent to which trained clinicians will make approximately the same judgement of a particular individual characteristic. In many areas of complementary medicine fundamental questions concerning the reliability of particular clinical judgements, diagnostic tools, patient classifications, or treatment effects have not been addressed by empirical research. On this basis alone, it is in many cases premature and misleading to draw too close analogies with orthodox medical research, where trials are often based on very specific research questions which have been shaped up by a wealth of accumulated data, and a coherent theoretical rationale.

Such work on the reliability and validity of basic clinical tools and techniques is an essential theoretical and empirical ground-clearing exercise, which needs to be carried out on a large scale to provide a meaningful context in which to undertake and interpret more sophisticated studies such as clinical trials. Unless we know that clinicians are talking about the same thing when they refer to, for example, particular homoeopathic types or cranial rhythms, studies based on such classifications will be very difficult to interpret. Similarly, it is essential to determine how much each of these classifications corresponds to clinically important differences between patients, which make a real difference when taken into account in treatment planning.

Such data is best obtained by trained clinicians carrying out studies to check the types of clinical judgement made in their own area of expertise. For example, Dove's*study of osteopathic assessments is an impressive example of a clear research question, i.e., 'can practitioners regularly and reliably agree with each other about clinical judgements they are making about a particular cranial rhythm?'. This was checked using simple, relevant and easily obtained information. Even apparently negative results showing low reliability can potentially lead to a re-shaping and refining of the research question; for example, is the inter-rater reliability of clinical judgements affected by the original training received by practitioners? or by type of cranial rhythm? or by transient patient factors? Even if subsequent studies show low reliability this is still useful in helping to stimulate the search for better diagnostic tools, and refine practice and training.

* Personal communicaton

Basic clinical questions

Kazdin (1980) has identified a range of questions which have provided a useful basis for psychotherapy outcome research, and which may also be of relevance to complementary medicine. I have elaborated on his ideas below.

Does treatment work?

The question here is whether patients change for the better as a result of treatment; would it make any difference if the patient was simply left alone? This simple question is often lost in a range of complex statistical manipulations yet it is the most fundamental issue in clinical research.

Patel (1987) has proposed that multidimensional cost benefit analysis is the most useful approach to treatment evaluation. This should include the cost of treatment required to produce particular therapeutic effects, in terms of time, money, other resources, and possible negative consequences for the patient. Clearly, all other things being equal, the least costly treatment alternative is to be preferred. Similarly, many different aspects of outcome should be taken into account in treatment evaluation. Patel suggests that holistic practitioners may see a higher proportion of positive treatment effects that are gradual and relatively intangible in nature, such as emotional and physical wellbeing, healthy lifestyle, and more satisfying relationships. Though such factors may be more difficult to assess than specific symptom reduction, they must be included in any comprehensive evaluation of treatment outcome and effectiveness.

Patient deterioration

Do some patients get worse as a result of treatment? This possibility is often obscured by the use of group averages which do not differentiate between those patients who do particularly well, and those who do particularly badly as a result of a particular intervention. There may also be a reluctance on the part of therapists of all persuasions to examine treatment failures. Though this tendency is understandable it may well be based on a fundamental misunderstanding.

Any effective intervention is likely to have the potential for causing negative as well as positive consequences depending on context. In the case of complementary medicine and psychotherapy, as with much of orthodox medicine, the question is to identify those applications which maximise positive and minimise negative consequences, rather than abandoning the treatment.The identification of deterioration in individual patients in response to treatment is an essential part of mapping out the optimal range of applications for particular treatments and specifying contraindications.

Durability

The temporal aspects of treatment and recovery are often overlooked in

comparisons of outcome in patient groups at a particular point in time. This approach is most useful in relation to acute conditions which can be described as 'cured' or otherwise following a limited period of treatment. For patients who do not fall into this category, closer examination of the nature of changes occurring over periods of time may help to clarify the processes involved in recovery. For example, some treatment approaches, such as homoeopathic medicine and psychoanalytic psychotherapy, predict temporary exacerbation of problems in some cases as part of the process of genuine recovery. These exacerbations are therefore not contraindications for treatment, but signs that it may be having a longer-term therapeutic effect.

Conversely, there is the important question of durability of treatment effects. There are real clinical and practical difficulties involved in producing and evaluating long-term effects, but they are an important aspect of comprehensive treatment evaluation. For example, research into the addictive disorders has shown relatively good short-term effects following treatment, followed by consistently high relapse rates within three months following treatment. It would be of interest to determine whether positive lifestyle changes are adopted as a result of complementary medical treatment, and if so whether they are equally ephemeral or more enduring.

Comparative effectiveness

The question here is which of a range of treatments is most effective for a particular condition; there are two related issues involved here. Firstly, there is a need for direct assessment of the nature and extent of effects produced by different treatments. Secondly, there is a need for conceptual clarification. Complementary medicine includes a wide diversity of disciplines, each of which share some basic common features such as concern with treating the whole person, as well as unique conceptualisations of human disease and distress, and associated interventions. Complementary disciplines are derived from diverse cultures and historical periods, often related to more primitive ideas and practices. An important part of comparative research must be the appraisal of the conceptualisations and procedures of each discipline, both in their own terms, and in relation to each other in order to provide a meaningful context within which comparative outcome studies can be designed and interpreted.

Causal mechanisms

What aspects of treatment are necessary, sufficient or facilitative of therapeutic change? This is a more focused question about the effects of particular components of the therapeutic package. It has attracted increasing attention in view of meta analyses of psychotherapy outcome studies, which tend to show fairly similar results despite differences in therapist, theoretical approach, and type of intervention. These findings strongly suggest that treatment may not always work for the theoretical reasons that therapists believe. Comparison of differential

outcomes following treatment does not by itself answer the question of what were the 'active ingredients' of treatment.

This question is usually addressed by identifying the main components of treatment, and assessing outcome in patient groups receiving systematically different combination of these components. For example, in homeopathy it would be possible to compare three groups of patients receiving only constitutional, only pathological, or both types of remedy. Similar studies could be carried out, including, for example, different lengths of consultation, advice on health maintenance, and physical examination, in combination with other treatment techniques. Often such research showing that particular aspects of treatment are associated with particular effects is accepted as adequate justification of clinical practice, though there may be considerable debate about the precise causal mechanisms involved.

Which therapist factors are important?

Within complementary medicine there has been much debate about the relative contribution of personal and professional qualifications to the therapeutic process. For example, homoeopathic medicine is practised by both medically qualified and lay clinicians, though the clinical importance of this difference remains unclear.

Psychotherapy research has shown that a wide range of practitioner qualities can affect the process and outcome of therapy including personality, interpersonal style, beliefs, gender, power, attractiveness, socioeconomic status, and length and type of professional experience. Of particular relevance is the fact that studies carried out in this area have shown little evidence that theoretical orientation of psychotherapists influences treatment outcome. On the basis of these findings, it could be hypothesised that diversity of theoretical orientations and techniques associated with complementary medicine obscures a greater similarity in types of personal qualities and approach relevant in the treatment process. The questions here are: what influence do the personal qualities of therapists have on treatment outcome, and do they bear any systematic relation to professional discipline or training?

Patient factors

As with therapists a wide range of patient characteristics have been shown to affect treatment outcome. One particularly interesting finding in psychotherapy research is that patients' perception of therapist qualities, specifically warmth, empathy and genuineness, are predictive of treatment success or failure in a variety of therapeutic contexts. By contrast, direct measures of therapist verbal and non-verbal behaviour ought to be expressive of these qualities, but they have not shown consistent links with outcome.

On the basis of such findings, Kendall and Norton-Ford (1982) proposed the hypothesis that patients' perceptions of therapy may have a more important influence on treatment than specific therapist actions or interventions. Put another

way, the patients' perception of therapy may be an important mechanism accounting for treatment effects.

Partly as a result of its search for objective and easily quantifiable effects, and partly due to the influence of therapist expectations, psychological research has only recently begun to take account of the patients' perspective. However, as Joynson (1977) has pointed out, one of the unique and most important characteristics of human beings is their capacity for self understanding and control. Any approach which tries to understand or predict human response without taking this important source of data into account is bound to meet with limited success.

Alternative strategies

Any research activity has two components which jointly contribute to defining the meaning of the research question. The *strategy* or design refers to the overall plan of campaign, according to which the study will be carried out and data collected. The *tactics* refer to more detailed methodology employed to achieve the wider objectives including how the researcher will interact with patients in order to study them, the precise data to be obtained, and the type of analysis to be carried out. Tactical issues are well documented in most research methodology texts, and will not be dealt with further here.

The clinical trial described above is only one of a range of potential research strategies, each of which has a different optimal range of application and different associated advantages and disadvantages. The following list is intended to give an idea of some alternatives to the clinical trial which have proved useful in psychotherapy research and may be of value to complementary medicine.

Single case design

It is often useful to ask what is happening to this patient? what am I doing to them? what experiences are they going through? how do they understand what is happening to them? are any events in their lives relevant to my intervention? The detailed analysis of individual cases can be used to illustrate certain principles, provide accounts of descriptive procedures, provide a focus for debate about the way in which particular problems can be managed, and generate hypotheses about significant factors in treatment response which could form the basis for subsequent studies using larger samples. This type of clinical vignette is commonly used for teaching purposes, though its potential in written form has been less well exploited.

This type of research may be particularly useful for holistic practitioners, who are concerned that the important role played by individual differences in treatment is obscured by use of standardised measuring techniques or average scores. It allows more detailed inclusion of the experience of both therapist and patient during treatment. A range of tactics can be used in the single case study, ranging from a narrative biographical format, to a controlled study in which the individual is used as their own control, and response compared at different stages of treatment, and

to different types of intervention.

The main limitation of this approach is that it may be unrepresentative and reflect the idiosyncrasies of a particular therapeutic intervention in a way which cannot be generalised to other contexts. This does not invalidate single case studies provided they are not used to address general questions such as the efficacy of particular treatments.

Questionnaires

Given the importance of subjective and experiential factors to the treatment process, the development of questionnaires to assess such material can potentially make an important contribution to research. However, they also introduce a range of new problems. It is a very common occurrence for researchers to mistake substantive questions about what they are intending to study for practical problems about the phrasing of questionnaire items. Partly for this reason, it is often preferable to find other ways of obtaining this data where possible, such as direct observations or measurements. William James, one of the founders of modern psychology, recognised this point as long ago as the 1890s in his statement, 'because of its ease of use, the questionnaire is the bane of modern society'.

The questionnaire is a complex and subtle instrument and nothing is more cost effective when you get it right, but getting it right is often a tedious process which involves precise definition of the focus of research interest, eliminating sources of potential ambiguity and bias, and a design which is clear, simple, easy to use and interesting to complete.

Field study

This involves careful analysis of an ongoing situation into which some kind of change is introduced; for example, changes in referral rates from other practitioners following provision of different types of information. The stress here is on examining naturally occurring changes and developments, rather than introducing changes for the purpose of research. The main advantage of this strategy is its focus on authentic real life events in the clinical situation. Conversely, the main disadvantage is the researchers' lack of direct control over the events being studied, so that it may be difficult to identify the nature and direction of causal relationships.

Laboratory model

This involves setting up experiments in carefully controlled situations, such as the clinical trial. The main question which this strategy is designed to address is the extent to which a particular causal agent is responsible for a particular effect. The limitations have been described above.

Consultancy activities

In this situation the question is 'how can I facilitate or improve a particular intervention process? What can I do to make it more effective?'. As with field studies, the main advantage of this approach is its close contact with real life situations where the researchers' main aim is to improve a particular situation.

De-constructing the therapeutic process

Throughout the consideration of the study of complementary research we regularly return to the need to understand exactly what is going on in any therapeutic practice. This, in many ways is the central question for research.

Most therapeutic practices are described in general, broad terms. Frequently some central procedure, whether it be manipulation or the insertion of needles or prescribing herbs or pills, is regarded as characterising the whole treatment process. Yet practitioners of all persuasions will insist that these procedures are only a small part of a broad therapeutic intervention. Research therefore must begin to unravel the different components of therapy, break it down into its critical constituents. In terms of a data matrix, the proposal here is to develop matrices in which the facets that make up the columns are drawn from all the stages of therapy - how the patient chooses a therapist, or has one chosen for him/her, the nature of the initial contact with the practitioner, the types of information elicited and so on, right through to the various types of therapeutic outcome.

Another elaboration of therapeutic practice requires that differences between practitioners are considered. In matrix terms this implies thinking of the rows as falling into sets according to the practitioner the patient in that row sees. In other words groups of patients could be examined in order to see whether the patterns in their area of the matrix differ from patterns in other areas. Provided the columns carry information on the practitioner, then further analysis could reveal differences between practitioners that are important for the patient's experiences and treatment outcome.

Developing an understanding of holistic systems

A further development of this central question is the explanation of the concepts used in complementary practice. In the discussions of complementary practices and their theories considerable obeisance is made to the concept of the patient as a whole person, but research activities do little to identify exactly what this entails or to set up studies that are consciously 'holistic'. The problem appears to be that little recognition is given to the assumptions about the nature of the whole patient, in making the relevant measurements. For example, when measuring the galvanic skin response the patient is dealt with as an electro-chemical organism. Identifying the angle of parts of the skeleton implicitly regards the patient as a mechanical device and so on. If a patient is to be treated as a whole then measurements should include

the mechanical, pharmacological, biological, psychological and social aspects.

In terms of the data matrix the facets that make up the columns should include examples drawn from all the different frames of reference for considering people. Of course, this is a demanding requirement to be applied to every single study. But at least the range of measures should be included in the full agenda for complementary research. An important emphasis here is that all the measures do have potentially equal weight. The idea that there are 'objective' measures that have a special place and supposedly 'subjective' ones, only of relevance for interest or background information, is completely at variance with the principles and theories articulated in most writings on complementary practices.

The central problem of reliability

So far it has been assumed that all the activities of medical practitioners are carried out with a crisp scientific precision, that the judgements made and diagnoses reached are incontrovertible amongst practitioners of a particular school of thought. Daily conversation suggests that this is not the case. Indeed an account of the many subtle cues which an acupuncturist or osteopath, for example, is called on to distinguish, leads to the conclusion that such skills are fruitfully open to scientific scepticism.

The establishment of relationships between facets across people requires that the values that are put into the columns of the matrix are robust and stable enough to allow important correspondences to emerge. In other words, if one of the columns was the practitioner's judgement of, say, strength of a particular pulse, but the practitioner was not exactly sure about that judgement, to the extent that the same pulse had an equal chance of being categorised as very strong or moderate, then categorisation of pulse strength may not be reliable enough itself to relate to any other measures made. All measurements potentially suffer from the problem of low reliability and so they must be developed so as to minimise the problem.

In essence, there are two related procedures for improving reliability of any categorisation scheme or measurement. The first is to be as clear and precise as possible about what exactly is being measured. In facet terms this means specifying as effectively as possible all the categories, the elements, of a facet. This greatly increases the likelihood that the person making the measurement (a patient by self-reporting, the practitioner by making an observation, or researcher using a measuring device) does indeed measure the same entity whenever it is present. It also increases the probability that two people will have a shared understanding of what they are measuring and record the same thing.

The second procedure is to make each measurement a number of times and in a number of different ways. Some summary or averaging of the measurements can then be made. This will increase reliability, provided all the measurements are of the same phenomena. This latter caveat turns out to be more problematic than might be immediately apparent, as will be shown in the following consideration of validity. But for the present the point that can be made is that it does not matter, and is actually helpful, if the data matrix contains columns that in various ways

repeat themselves. In the initial stages of research there is some value in having redundant facets. These help to establish the basic reliability of the measurements. The relationships between similar measures not only helps to clarify what is being measured, but also helps to establish what the minimal reliability of the measurements are likely to be.

Validity and sensitivity

The consistency with which any measurement is made, usually discussed under the label of 'reliability', is distinguished in the research methodology literature from 'validity'. This is taken to be the extent to which a measurement measures what it is purported to measure. A measurement that was supposed to assess muscle tone, but actually measured blood pressure, would he quite misleading, no matter how reliable it was. The problem is, though, establishing what any system of categories of numbers is actually measuring. Pupil dilation, for example, has been discussed in the popular press as an indication of interest or excitement, but it may be more of a general autonomic arousal response that can be induced by fear, anger, joy or whatever.

The only way to establish what a measurement does reflect is by a scientific process. This requires understanding as to what other measurements it should relate to according to the theory and then demonstrating that data show those relationships do hold. In this framework all studies using the various measurements are part of the process of validating those measurements. The relationship between the facets that is explained by the theory is a test of the validity of those facets. From a facet perspective the establishment of a correspondence between some aspect of the definitional system and the observations plays two parts; it contributes to establishing the validity of the theory and also to demonstrating the validity of the method of making the observations.

Reconsidering the problem of reliability, it will be recalled that this derived from establishing the relationship between clearly defined similar measures. But validity has just been shown to be about such relationships as well! The inevitable conclusion is that validity and reliability are both part of the same scientific process. Their difference is one of emphasis. In the early stages of a study it might be questioned as to whether people who declare themselves as worthless, say, do also consistently reveal themselves as unprepared to face challenges. Some psychological theory may suggest that a measure of one should relate to a measure of the other. If such relationships are found then their validity is enhanced. But then both may be taken as indicators of self esteem, their relationship being an indication of the reliability of that measurement.

This discussion may be taken one important step further. A very precise, closely defined measuring procedure must still be shown to relate to some other procedures for its reliability and validity to be established. But consider, for example, the measure of blood pressure as a general indication of the stressful state of the individual. Some of the unreliability of this indicator of stress comes from the relationship of blood pressure to age, to recent physical exertion and many other

aspects of the individual. It is sensitive to other things than just stress. However, if it were only related to stress (assuming there is a clear definition of that) then it would he of little value. It would not be possible to relate stress to anything else *but* blood pressure.

The point here, and it is critical for all research in complementary medicine, is that a measurement can be *too* reliable, its sensitivities can be too focused. All scientific activities grow out of the establishment of relationships between different facets, continually generating superordinate facets that encompass previously distinct ones.

Understanding the role of patient variation

All the matrixes and facets discussed so far have focused on groups of patients and comparisons between them. However, much of the discussion of the differences between people and significance of that difference for treatment is really derived from the notion that comparisons between people even on a number of facets misses the important qualities of their essential uniqueness. This idea can be given a much clearer scientific definition by casting a matrix for each individual. By treating the same person at different points in time as the rows, or by taking different aspects of their problem as the rows, a variety of facets can be generated. This will enable the particular structure of the individual's illness to be examined. The special quality of that person's illness can then be explored, but comparisons with other patients is not ruled out. The comparisons can be made of the internal, numerically established patterns. This will give a comparison of the structure of the relationships between the different problems as manifested in different patients.

The psychological literature is full of such studies and many of the questions raised by Heron and Reason (1986) in the first issue of *Complementary Medical Research* were concerned with developing parallel studies in the medical arena.

Hypothesis generation as well as testing

When the nature of the research activity is discussed at the level of abstraction and in the degree of detail presented here the novice and expert are both likely to express horror at its complexity and the difficulty of knowing where to begin. There is certainly little doubt that discovering fruitful starting points are some of the most valuable contributions of any research activity. The discovery of the structure of DNA did not, of itself, have immediate practical implications, but it opened the doors to many others.

One of the objectives of complementary medical research should therefore be to find effective future starting points, i.e., finding ways of thinking about problems. In the language of science it means carrying out studies that will generate hypotheses as well as studies that test them.

The idea of hypothesis generation research can be thought of by considering as an example a study in which a number of patients have been treated by a variety of practitioners and the effects of that treatment have been recorded in a number of

different ways. If some means could be found of representing all those patients so that previously unthought-of aspects of the differences between them could be related to their treatment experience, and brought to light, then this might suggest directions for more detailed research.

Having attention drawn to previously unconsidered aspects of people may seem a very demanding procedure, but once again the faceted data matrix helps to make the task manageable. Instead of starting with a set of facets and seeing whether the patterns in the data matrix support the validity of those facets, essentially a hypothesis testing procedure, the reverse is proposed. The matrix is examined to see if it reveals any interesting patterns and the researcher/practitioner is then challenged to see if those patterns suggest any possible explanations, generate any hypotheses. Computer based procedures now exist that help to indicate which patterns in the data may be of interest.

Strategies for context manipulation

Like a war on ignorance, research can be thought of in terms of the overall plan of campaign, the strategy, and in terms of the particular ways in which battles are fought, the tactics. A full range of both strategies and tactics must be drawn upon to have some hope of making progress in the war. The variety of tactics available has already been discussed in terms of the measurements to be made, the facets that define the columns of the data matrix. It has also been emphasised that any such research tactic carries with it assumptions about the type of organism that is being studied. A truly holistic approach therefore requires that the full range of tactics be employed.

By contrast, the strategies used make assumptions about the types of research questions being asked. In terms of the data matrix, research strategies are ways of thinking about the analyses to be performed on that matrix, or the types of patterns in that matrix that will be of significance. For example, if the efficacy of a known causal agent, for instance a drug, is in question then the researcher needs to be confident that s/he has direct control over who receives that particular agent and who does not. The only way to do this is for treatment agents to be randomly distributed among a reasonably homogeneous set of patients. This is the randomised experiment. Clear differences between measurements in one set of rows when compared with another set in the data matrix will be evidence for the effects hypothesised. As can be seen, the strategy is concerned with what aspects of the context under study need to be manipulated in order to answer a particular type of research question.

Often the precise question of causality is not at issue. Rather the researcher wishes to know whether there is convincing evidence that some approach to treatment has recognisable consequences. In this case, rather like the legal attempt to provide proof of guilt, the researcher will attempt to amass information from different sources, comparing various possibilities, attempting to corroborate and validate the effects of the treatment. Rather than answering *causal* questions as in

the experiment, the attempt in this strategy is to answer *probitive* questions.

Beyond the questions of treatments and their outcomes there are a variety of questions, possibly the majority in the medical area, that are really concerned with establishing the pattern of relationships between the variables that exist, This third, relational, strategy is the one drawn upon most usually by epidemiological studies. Here the matrix is examined for the pattern of relationships it reveals and possibly for changes in relationships over time. Frequently this can be a hypothesis-generating procedure.

A fourth strategy, already mentioned in relation to research objectives, is important to record. This is the comparative case study. In many areas of medical activity it is very important to try and establish what exactly is going on in specific cases. Here the broad relationships across individuals is less a question than the particularities of individual cases. In the history of medicine such detailed studies of individuals have often been of great significance. There are now detailed guidelines available for the conduct of case studies beyond the conventional medical context (Bromley, 1986). Any research programme may well move back and forwards between different strategies as the research questions being asked change and unfold.

Recognising the restrictions on research resources

The discussion above has thrown a broad net over many possibilities. Clearly resources are limited for doing all this work. It is therefore worth, finally, considering the nature of these limitations and the opportunities available for overcoming them.

The major limitations may be listed as follows:

(a) Readily available, reliable and robust data collecting procedures.
(b) The data collecting and analysis skills necessary to use these procedures.
(c) The opportunities for collecting and analysing information about patients and treatments.
(d) The development of clear, testable accounts of the principles involved in complementary practice.
(e) The difficulties inherent in obtaining the patients' experience of complementary practices.

A broader view of science

One major problem with the scientist-practitioner model has been identified as excessive reliance on strict experimental methodology in the clinical context. The history of science shows that this approach of transferring methodology from one field to another is rarely productive, and genuine progress is more likely when attention is first centred on the questions which need to be asked; only then can an

appropriate methodology usefully be identified. In support of this suggestion, some commentators have noted that social work seems to have produced more relevant and applicable research than clinical psychology. This has been attributed to the fact that social workers, unlike clinical psychologists, do not receive a prolonged training in experimental research methodology. As a result, more emphasis is placed on asking the right questions first, and then using whatever methods seem most appropriate to finding the answer.

One of the motivations underlying the tenacious hold of experimental methodology on clinical research has been the legitimate desire to carry out genuinely scientific work. The problem here is to define what 'science' really means. It has long been recognised that science does not mean building grand metaphysical systems to explain the world; that is, or at least was, philosophy. More recently, it has become apparent that science is not about collecting all the information and ending up with huge piles of data, or following a particular methodology, or using complex statistical techniques either. Rather the discipline of science consists of looking at the correspondence between the research question and the empirical data. This depends on establishing a clear question, and a methodology for obtaining data of direct relevance to the question. Data can then be used to further refine the question, leading to genuinely cumulative development in empirically-based knowledge. It is the 'goodness of fit' between the question and the data that forms the basis of good scientific research.

One example of how empirical observations can lead to clearer conceptual classification schemes is the discovery that some types of material leave marks on photographic plates, which led to the classification of materials in terms of whether or not they emit radiation which can be recorded. This way of thinking implies that there is some conceptual way of dividing up whatever it is you are studying, such as remedy types, acupuncture points or types of illness. The observations which are made, and the data obtained are ways of classifying the answers. Statistics is of course concerned with technical ways of classifying the data, and should ideally be used to assess the degree of correspondence between the conceptual divisions made in the research question, and the patterns which emerge from the data.

It is in the nature of science always to have a balance between precise confidence in any key concept and sceptical questioning of exactly what is the focus of concern. Research should therefore be seen as an unfolding, clarifying and defining activity that has no obvious end point. There is no logic to looking for *the* cure for a given illness. The research activity is concerned with an ever-greater elaboration of what is the nature of the cure and the nature of the illness and the relationships between these two. On the strictest scientific grounds the study of complementary medicine is essential for the scientific development of orthodox medicine.

References

Barlow, D.H., Hayes, S.C. & Nelson, R.O. (1984), *The Scientist Practitioner,* Pergamon, Oxford, pp. 3-33.

Bergin, A. & Strupp, H. (1972), *Changing Frontiers in the Science of Psychotherapy*, Aldine, Chicago.

Bromley, D.B. (1986), *The Case-study Method in Psychology and Related Disciplines*, Wiley, Chichester.

Canter, D. (1985), *Facet Theroy*, Springer-Verlag, New York.

Comber, M. & Canter, D. (1983), 'Differentiation of Malicious and Non-malicious Fire Alarm Calls using Multidimensional Scaling', *Perceptual Motor Skills*, Vol. 57, pp. 460-2.

Fulder, S. (1984), *The Handbook of Complementary Medicine*, Coronet, London.

Garfield, S.L. (1981), 'Evaluating the Psychotherapies', *Behaviour Therapy*, Vol. 12, pp. 295-308.

Gough, G. (1985),' Reasons for Slimming and Weight Loss', in Canter, D. (1985), *Facet Theory*, Springer-Verlag, New York, pp. 245-56.

Henkel, R.R. & Morrison, D.E. (1970), *The Significance Test Controversy*, Butterworths, London.

Heron, J. (1986), 'Critique of Conventional Research Methodology', *Complementary Medicine*, Vol. 1, pp. 12-22.

Joynson, R.B. (1977), *Psychology and Common Sense*, Routledge and Kegan Paul, chapter 1.

Kazdin, A.E. (1980), *Research Design in Clinical Psychology*, Harper and Row, New York.

Kendall, P.C. & Norton-Ford, J.D. (1982), 'Therapy Outcome Research Methods', *Handbook of Search Methods in Clinical Psychology*, by Kendall, P.C. and Butcher, J.N., Wiley, New York.

Kiesler, D.J. (1965), 'Some Myths of Psychotherapy Research and the Search for a Paradigm', *Psychological Bulletin*, Vol. 65, pp. 110-136.

Levy, S. & Guttman, L. (1985), 'The Partial Order of Severity of Thyroid Cancer with the Prognosis of Survival', *Data Analysis in Real Life Environment* by Marcotorchino, J.F. et al, North Holland, Amsterdam, pp. 111-12.

Orne, M. (1962), 'On the Social Psychology of the Psychology Experiment with Particular Reference to Demand Characteristics and their Implications', *American Psychologist*, Vol. 17, pp. 776-83.

Patel, M.S. (1987), 'Evaluation of Holistic Medicine', *Social Scientific Medicine*, Vol. 24(2), pp. 169-75.

Raimy, V.C. (1950), *Training in Clinical Psychology* (Boulder Conference), Prentice Hall, New York.

Strupp, H.H. (1981), 'Clinical Research, Practice and the Crisis of Confidence', *Clinical Psychologist*, Vol. 49, pp. 216-20.

Tyerman, C. & Canter, D. (1983), 'A Taxonomy of Small Hostels for Alcoholics', *Drug and Alcohol Dependence*, Vol. 11, pp. 225-31.

Wharton, R.L. & Lewith, G.T. (1986), *Complementary Medicine and the General Practitioner*, report to the RCCM.