2 Some basic problems in evaluation

Summary

In this chapter we review some of the ways in which investigators have tackled fundamental problems pertaining to the complex task of evaluation of psychotherapy.

The three basic components of psychotherapy have been described as (a) the patient and his or her problems; (b) the therapist, his or her personality, style and technique; (c) the period of therapy and its aftermath. To these we added a fourth – the psychosocial environment in which treatment takes place. In the following pages we examine each of these components and their impact on the treatment of maladjustment in children in ordinary junior and senior schools.

Under the heading 'children and their problems' we discuss the following in relation to treatment: the severity and type of disorder; educational issues; and the age and developmental stage of the child.

We consider the controversy about the relative importance of the therapy technique and the child's personal relationship with the therapist. We discuss therapists' styles (directiveness, warmth, empathy, genuineness, activity level) and the effect these qualities may have on the child. Evidence is presented that points to other people, such as teachers, parents, nursery nurses, and teacher-aides, being able to make a worthwhile contribution to treatment.

With regard to the effect of time in psychotherapy, we stress the need for protracted follow-up studies after treatment has ended, and show that their timing may be crucial. We present the rationale for short-term therapy, in which it is argued that brief and intensive treatment may be more effective than long-term approaches.

The psychosocial environment (both home and school) has a

profound effect on a child's behaviour and development, and thus on any psychotherapeutic regime. We consider that it may be one of the factors influencing spontaneous remission.

We draw attention to the different types of control groups, and the importance of assessing whether or not the members of these groups are receiving some form of 'treatment'. We also examine the concept of spontaneous remission.

Introduction

Whatever his or her origin, be it the field of practical psychotherapy or academic psychology, the newcomer to psychotherapy research is likely to be overwhelmed by the complexity of the task of evaluation as there seem to be so many subtle and shifting variables that may be important, and so many conceptual as well as practical logistic problems to overcome. This chapter selectively reviews some of the ways in which psychotherapy research workers have tackled some of these basic problems.

The early approach, in the USA, used by Eysenck (1952) for adults and by Levitt (1957) with children, was simply to ask, 'Is psychotherapy effective?' Both authors pooled the results of previous studies and came to the conclusion, now much discussed and criticized, that there was no evidence that psychotherapy was effective. We now examine in some detail the criticisms that have been made of these reviews, particularly of Levitt's work because this concerns children.

One important criticism is that the question asked by these early workers was too general. As outlined at the start of the chapter, psychotherapy has three basic ingredients, each of which is almost infinitely variable and must be considered when developing and testing theories of psychotherapy (Kiesler 1971). To reiterate, these ingredients are (a) the patient and his or her problems; (b) the therapist, his or her personality, style, and technique; and (c) the period of time over which the therapy occurs and over which its effects may become manifest. Our community-based approach led us to add a fourth ingredient, the psychosocial environment in which treatment takes place. Levitt (1957) was, in fact, aware of the need to differentiate some of these ingredients. He tried in his review to include only those studies in which the patient population suffered from emotional and neurotic problems and he distinguished between measurements made at the end of treatment and after a follow-up period. However, the types of treatment used in the studies he pooled were heterogeneous, and the suggestion that he added together the results of quite dissimilar studies is further supported by

the fact that *within* his treatment group the reported outcomes were quite different.

A more sophisticated method has been used with adults. This consisted of combining the results of different studies, a task undertaken by Glass and Smith (1976) and Smith and Glass (1977) in a survey of 375 controlled studies with differing patient groups, treatments, and follow-ups. These authors calculated the results of each study in a way that was statistically comparable with the others (they calculated the mean difference on outcome measures between control and treatment groups divided by the standard deviation of the control group). They were able to show that the 'average' client receiving therapy was better off than 75 per cent of the untreated controls. It would be interesting to repeat this analysis for the results of child psychotherapy studies, as it failed to support Eysenck's pessimistic conclusions. We return to Smith and Glass's work in Chapter 10. Meanwhile, as both Gottman and Markman (1978) and Kiesler (1971) pointed out, we are more likely to gain an understanding of psychotherapy by looking at its key ingredients separately rather than in combination with one another. We will now look at these components in terms of their affect on our own research.

The children and their problems

The first ingredient mentioned by Kiesler is the characteristics of the patients and their problems.

THE SEVERITY OF DISORDER

We defined a disturbed child as one in whom there was some *demonstrable* abnormality of behaviour, emotions, or social relationships. The simplest dichotomy of characteristics was, therefore, between disturbed and non-disturbed children. Caplan (1964) has suggested that all disorders are present at the outset in an undiagnosable form and that they develop from this into manifest disorders. In his concept of 'secondary prevention', he argued that it may be advantageous to take action very early in the development of a disorder. In our community study we were in a good position to examine Caplan's concepts.

The first difficulty we faced was that of identification. Escalona (1974), in a discussion of intervention programmes for children at psychiatric risk, pointed out that there is a range of minor neurological abnormalities and, particularly, social adversities which is associated with a raised incidence of psychiatric disorder and, as such, can be seen as a risk factor. The problem is that, unless we

undertake generalized and massive efforts to alter the life experience of young children, it is only when minimal but overt deviations of development occur that any specific interventions become possible. Among young children who are in the process of development it seems important to offer help to a wider group than just those with established disorders. The detection of less severe abnormalities seemed to us to offer a way of doing this. We postulated that children with both established and less severe disorders are 'at risk' in terms of the extension or progression in severity of such difficulties (Kolvin *et al.* 1977). All these children should therefore be offered help early on. Our view was that there was already evidence of an association between educational, emotional, and relationship difficulties, and that study of the processes of treatment in an at-risk group comprised of children with both established and less severe disorders would help us to decide which children respond best to intervention.

There is another sense, however, in which treating mild disorders is important. Because some of the children we see have such serious handicaps, we assume that these cases must have absolute priority in treatment. On the other hand, there seems little point in mounting complex treatment unless the treatment is effective. We also need to study the impact of treatment on milder disorders, because it is important to know whether treatment is effective for them.

THE TYPE OF DISORDER

We have referred to the need to consider the type of disordered behaviour that a child is showing. The type of disorder affects the choice of treatment because there is no reason to think that the same therapy will be effective for all types of disorder. As the field is so large, we will give only a brief didactic account of the evidence of effectiveness of psychological and educational treatments on various types of disorders.

Neurotic disorders

Although we would have expected numerous attempts to evaluate the effectiveness of psychotherapy in neurotic disorders, the available literature was, in fact, very sparse. In a series of studies, Eisenberg and his colleagues (Cytryn, Gilbert and Eisenberg 1960; Eisenberg *et al.* 1961) compared the effects of drugs and psychotherapy on various diagnostic groups. As might be expected, judging from the results of other follow-up studies, the neurotic-disordered group showed greater improvement under both treatment conditions than did the conduct-disorderd group, but psychotherapy showed no advantage over drug treatment. It should be noted, though, that the psycho-

therapy given in this study was rather ill-defined and attenuated.

One major study on evaluation of treatment of phobic children was the important work of Miller *et al.* (1972). This study involved random allocation to treatment and control groups using time-limited treatment (eight weeks at three sessions per week) and same-sexed therapists; ratings at pre-, post-treatment, and follow-up (fourteen weeks, one, and two years) undertaken by independent raters; control of initial values by covariance; and the use of multiple outcome measures. The two treatments studied were psychotherapy and systematic desensitization. Only forty-four children were treated and there were twenty-three waiting-list controls. The results of the treatments did not differ, but both differed from the results of the waiting-list controls, on the basis both of the evaluating clinician's ratings of severity and of behaviour scales completed by the parents. Follow-up two years later showed that children who were successes at the end of treatment continued to be successes. The rates of improvement were more impressive for younger than for older children, as shown in *Table 2(1)*. However, sex, IQ, socioeconomic status, and chronicity were not related to outcome. The only other relevant finding was that children of highly motivated parents were more likely to succeed.

Table 2(1) *Treatment of phobic children* (from Miller *et al.* 1972)

	under ten years				*over ten years*				*total*			
	treated		*controls*		*treated*		*controls*		*treated*		*controls*	
	n	%	n	%	n	%	n	%	n	%	n	%
successful	23	96	8	57	9	45	4	45	32	73	12	52
unsuccessful	1		6		11		5		12		11	

Conduct disorders

These problems give rise to more social distress and concern than neurotic disorders and carry a worse prognosis (Robins 1966). A review of relevant treatment studies had to include some studies of delinquency, as the chronic offender is very likely to fall into the category of conduct disorder. Many of the studies, both of counselling (Powers and Witmer 1951) and residential programmes (Clarke and Cornish 1977), yielded negative results. Nevertheless, the studies described below yielded positive results and may augur well for future work in this area.

Adams (1970) evaluated a counselling programme for delinquents

in a custodial setting (the PICO Project). The study compared the reconviction rates for two sorts of offender (both of whom had a counselling experience) with those of comparable control groups. The first type of offender showed awareness of problems, verbal ability, high anxiety, and a desire to change; offenders in this group were termed 'amenables'. The second type ('non-amenables') did not show these characteristics. After release on parole the amount of time spent in any kind of custody was recorded. It was found that the non-treatment controls of both 'amenables' and 'non-amenables' had spent a similar time in custody, whereas the treated 'amenables' had fared considerably better. The treated 'non-amenables' had done worst of all. This study illustrated to us the importance of examining separately children with different characteristics.

Some institutional programmes have operated on the basis of psychodynamic theory. For example, at Northways House (Miller 1964) some severely disadvantaged delinquents were selected, following Borstal training, and were offered a residential treatment. The boys were expected to go out to work and the establishment was run along the lines of a normal residence for boys of their social group, with the inclusion of psychoanalytic components. Although numbers were small, there were encouraging signs, when compared with a control group, that the Northways boys were functioning better at follow-up in that they had lower reconviction rates, more settled marriages, and better employment records than control boys with similar Borstal records and background characteristics.

More recently, encouraging findings have been published on the effectiveness of behaviour modification in conduct disorders. Patterson and his colleagues (Patterson, Cobb, and Ray 1973; Patterson 1974) developed a home-based treatment programme for children with problems of conduct. This complex, four-staged programme involved instruction to the parents about behaviour modification principles, followed by analysis and modification of behaviour. These workers have reported very significant reductions in the rates of particular types of rowdy behaviour over the course of treatment. Alexander and Parsons (1973) compared family behavioural treatment of delinquents with other approaches and reported that the behavioural regime was the only one that led to a significant reduction in recidivism.

Behaviour modification techniques have also been used with notable success in residential settings. In the Achievement Place Project the assumption was that the boys lacked social skills. These were inculcated by a comprehensive curriculum of training in such skills, which was linked with a token economy system (tangible

rewards for achievement). There was a self-government system and relationships with staff were informal (Willner *et al.* 1978). Objective evaluation has shown this to be an effective form of management.

Table 2(2) *Outcome of treatment in different categories of psychiatric disorder* (Barrett, Hampe, and Miller 1978, from data by Levitt 1963)

improvement	*neurosis* n = 230	*special symptoms* n = 213	*mixed* n = 697	*acting out* n = 349	*psychosis* n = 252	*total* n = 1741
	%	%	%	%	%	%
much	15	54	20	31	25	26
partial	46	23	48	24	40	39
none	39	23	32	45	35	35

Barrett, Hampe, and Miller (1978) reorganized Levitt's data (which relates to the USA) to compare the progress of different categories of disorder. *Table 2(2)* is an adaptation from the 1978 work and shows the different outcomes. The neuroses in this table appeared to be the type more commonly found in adults than children and, therefore, were likely to be of a deeply ingrained or intractable type, with only a small percentage showing impressive improvement. The mixed category, which apparently consisted of general child guidance cases, gave rise to a similar picture. On the other hand, it was not unexpected that the category of special symptoms, which included a number of conditions known to improve spontaneously with time, such as enuresis, tics, and phobias, showed impressive results. The acting-out category, which was likely to contain a large number of children with antisocial behaviour, had a high percentage of cases that showed no improvement. By British criteria, it is unlikely that one in seven cases treated in child psychiatry would be considered psychotic – nor indeed would such high rates of improvement be expected with psychoses. We therefore suspected that these cases would be placed in a mixed category in the UK.

Educational problems

Many approaches to educational problems involve components that might reasonably be called psychotherapeutic. Two examples are early intervention (Head Start) programmes and special educational settings (Project Re-Ed). There is likely to be an overlap between educational problems and conduct disorders and many of the projects were designed to deal with both.

The Head Start projects have been reviewed by Bronfenbrenner

(1974). This is a most important review because, instead of giving a depressing 'thumbs down' to the whole Head Start movement, Bronfenbrenner has tried to co-ordinate aspects of the various evaluation projects that did give positive results and to integrate them into a picture of what could constitute success. The picture he presented was briefly as follows:

(i) When compensatory stimulation is provided for the pre-school child there are substantial IQ gains while the programme lasts but, after a year, this trend reaches a plateau, with gains becoming rapidly eroded once the help ends (DiLorenzo 1969; Gray and Klaus 1970; Weikart, Deloria, and Lawson 1974). Deutch (1971) saw this erosion of gains of an enrichment programme, even when it was still continuing, as being determined by social and family factors beyond the school, citing the fact that the children whose response was poorest came from the poorest environments.

(ii) The hope that programmes started early in life would produce the greatest and most enduring gains has not been fully substantiated (Braun and Caldwell 1973). Children involved in pre-school programmes, joined before the age of three years and *not directly involving mothers*, did no better than those who entered later, although their programmes were of equal duration.

(iii) Hays and Grether (1969) found that the lack of stimulation experienced by the disadvantaged child over the long summer holidays appeared to be responsible for much of the loss incurred in the areas where gains had previously been made. This is a strong argument in favour of home-based programmes which are, obviously, not subject to this problem.

(iv) Indeed, home-based intervention has led to dramatic and enduring gains, three or four years after help was stopped. There is, however, one important qualification – maternal interest and participation in the scheme are essential, as shown in (ii) above. A one-to-one interaction between a motivated mother and her child, around a common educational-type activity, was found to be crucial (Levenstein 1970; Schaefer and Aaronson 1972). The earlier the interaction began, the greater the gains appeared to be Karnes *et al.* 1968; Gilmer, Miller, and Gray 1970; Levenstein 1970): they were negligible if interaction started late. It is essential that such home-based educational programmes be reinforced when the child's dependency on his or her mother is greatest, that is, in the second year of life (Bronfenbrenner 1968; Levenstein 1970). So impressive were these findings that Radin (1972) suggested parent education is an essential adjunct to any compensatory pre-

school programme if the child is to continue to benefit cognitively.

(v) The optimal time for parental involvement seems to be in the first three years of the child's life. Nevertheless, there is considerable evidence to show that parental involvement thereafter continues to be an important factor in the educational progress of the child (Smith 1968). Some families (especially psychologically vulnerable families) are so socially disorganized that the parents are unlikely to be able to participate. In such circumstances a more radical solution has been attempted – the separation of infant and mother during the waking day, the provision of compensatory stimulation for infants (Heber *et al.* 1972), and the training of mothers in child-rearing and basic domestic skills. Such radical intervention would appear to be intellectually, educationally, and morally justifiable only when home conditions appear to be totally detrimental to child development. However, not only are the costs of such projects prohibitive but, of equal importance, we know little about their social and emotional consequences.

At least four important conclusions relevant to preventive child psychiatry have emerged from educational research. First, with disadvantaged children greater educational gains are likely to result from more, rather than less, structured educational programmes. Second, there seems to be an overriding need for the involvement of the mother and child in a common educational task. Third, erosion of gains after help has stopped must be recognized as a problem. Finally, Bronfenbrenner (1974) stressed the need to improve the total living conditions of disadvantaged families in the community. It is unfortunate that the Head Start projects concentrated almost entirely on educational and cognitive measures. Behaviour ratings were often made, but were generally so diverse that no general conclusions could be drawn.

Other educational evaluations have been linked to special educational projects of various sorts. Quay *et al.* (1972) reported the effects of a part-time resource room run within the ordinary school. After one year of this special regime the experimental group had made significant gains in academic achievement. There were very marked differences between the children's behaviour while in the special and then ordinary classes, measured concurrently. Another widely known educational treatment is Project Re-Ed (Weinstein 1969). This consisted of a short-term residential educational treatment programme. Measures were made before and after the programme by the referring school, parental reports, and achievement tests. In the early reports there was, unfortunately, no control group, but it is

interesting that the most spectacular improvements were those reported by parents.

Neurotic disorders, conduct disorders, and educational problems constitute the three most relevant types of problem that children are likely to have. There are, however, many other important ways in which children may differ.

THE AGE AND DEVELOPMENTAL STAGE OF THE CHILD

It is an index of the early and primitive state of child psychotherapy research that so few studies have seriously looked at the effect of psychotherapy on children of different ages. An exception is the Miller (1972) study quoted above (p. 20). This demonstrated that not only was the outcome of phobic states better in younger than older children, but also that treatment was more effective amongst the former.

The therapist: his or her personality, style, and technique

There is some controversy as to whether or not the therapist is more important than the therapeutic technique.

THERAPISTS' STYLE

A great deal has been written about therapists' styles and their relationship to other factors in therapy. It seemed obvious to early researchers (e.g. Fiedler 1950) that practitioners from different theoretical schools of psychotherapy would show differences in the way that they approached their patients. Different theoretical schools do, indeed, suggest very different approaches to the patient. For example, Rogers's (1952, 1959) non-directive therapy is extremely brief and passive compared with the prolonged treatment recommended by psychoanalysis. Sullivan's (1953, 1956) technique, influenced as it is by the view that the self is made up of others' views of the individual, might be expected to be very different from the psychoanalysts' view that biological drives are of fundamental importance in personality development. It must have come as a surprise, then, to these early workers to find that the theoretical school of the practitioner was fairly unimportant compared, for example, with the differences between inexperienced and experienced therapists from the same school (Luborsky *et al.* 1971, 1975). Strupp (1958) carried out detailed studies of interaction in psychoanalytic and Rogerian therapies. He confirmed that differences existed between the approaches, and that these were very much in line with what might be expected on theoretical grounds. More recent work has been reviewed

by Pope (1977). He concluded that theoretical orientation is one of several influences on therapist style. In factor analytic studies, various main characteristics of style have been identified: for example, nurturance, a non-judgemental approach, an analytic approach, etc. Little has been done to relate these to reactions in the client or indeed to the outcome of therapy. Unfortunately, no one has examined how a particular treatment style is modified by different treatment settings, for example, psychoanalytic therapy applied to groups compared with casework. It is just such comparisons that would have been of the greatest relevance to the present research.

Therapist directiveness

The first of the more specific elements of the therapist style is therapist directiveness. This is defined slightly differently from study to study, but generally consists of the extent to which the therapist leads the interaction and is prepared to introduce themes or to make interpretations ahead of the client and in areas of which the client is unaware. It also includes advice-giving. Ashby *et al.* (1957) undertook a study with six therapists, three of whom were trained to be 'directive' and three 'reflective' therapists. Psycho-neurotic patients were randomly assigned to the therapists. The most striking result of this study was that certain types of patient seemed to respond better to the directive therapists, while other types preferred the reflective therapists. Those patients who had been the most defensive before treatment behaved more defensively with a leading (directive) therapist than did those who manifested a need for autonomy before treatment started. The latter group reported feeling less defensive during directive therapy than during reflective treatment.

More recent studies have confirmed this differential effectiveness of directive and reflective therapist with different types of client. For example, in two studies (Abramowitz *et al.* 1974; Friedman and Dies 1974) clients were assessed as internally directed or externally controlled on a locus of control categorization (Rotter 1966). As expected, the externally controlled group were best with directive therapy and the internally controlled one preferred the non-directive treatment.

Warmth, empathy, and genuineness

These are the most intensively studied therapist qualities, and several extensive reviews have been produced. Truax and Carkhuff (1967) claimed unequivocally that these three qualities, which may be measured from brief tape-recorded segments of therapy sessions, were necessary for the success of therapy with a wide range of client groups. There has been a great deal of discussion as to what the

qualities actually constitute – particularly because accurate empathy is concerned with unobservable phenomena. A more recent review (Mitchell, Bozarth, and Krauft 1977) acknowledged that results of the early studies had not always been replicated. With different types of therapies and different client groups there may be major variations in the importance of these qualities.

Another point that gives rise to uncertainty is the absolute level of the therapist qualities. Mitchell, Bozarth, and Krauft (1977) commented that if the therapists are generally below an absolute cut-off level of empathy, warmth, and genuineness the conditions fail to operate and no correlation with outcome can be expected.

Considering the importance attached to these qualities in research with adults, it is surprising that they have not been studied to a larger extent in research into child psychotherapy. Siegel (1972) found that improvement among children with learning disabilities in both verbal and behavioural spheres related to length of time in play therapy and to therapist levels of empathy, warmth, and genuineness. Truax *et al.* (1973) found more mixed results. However, parents' perceptions of improvement were related to the therapeutic qualities of the therapists.

Therapist activity level

This has been studied by Lennard and Bernstein (1960) and very extensively by Matarazzo and colleagues (Matarazzo *et al.* 1965; Matarazzo *et al.* 1968). Interviewer activity in these studies was measured by the *amount* of therapist or interviewer activity, regardless of its *content*. There is no evidence that therapist activity measured in this way directly affects outcome: however, an active therapist is more likely to be seen by the patient as warm and be able to put the patient at his or her ease than is a more detached therapist. In one study there were significantly fewer missed appointments with active therapists (Lennard and Bernstein 1960). Both warmth and patient attendance are important variables in the success of therapy.

THE EFFECTIVENESS OF NON-SPECIALIST HELP

Adams (1975) wrote disparagingly of the extent to which children are treated, not directly by highly trained mental health professionals as adults might be, but by other workers such as teachers, parents, nursery nurses, and teacher-aides. Nevertheless, there is evidence that such people can make a considerable contribution to treatment projects. Carkhuff has undertaken a number of studies and has reviewed the field (Carkhuff 1968; Anthony and Carkhuff 1977). Three studies in the mental health area suggested that non-professionals can be as effective as fully qualified professional

workers (Poser 1966; Zunker and Brown 1966; Truax and Lister 1970). While there have been some criticisms of the methodology of these studies – particularly the dissimilarity of the professionals and non-professionals on criteria other than that of their qualifications – there is no evidence to contradict these findings.

In the educational field, both teachers (Aspy and Roebuck 1971) and teacher-aides (Cowen *et al.* 1975) have been used in treatment projects. Again, there is evidence of the positive impact on student mental health of the teachers' skill in human relationships. These results are discussed more fully in Chapter 6.

The dimension of time in psychotherapy

The effects of psychotherapy may increase or decrease with time, they may be seen only after a latent period (even after treatment has stopped), and may cease altogether after a while. It is therefore very important to have short, intermediate, and long-term follow-up studies.

DURATION OF TREATMENT

Strupp (1978) offered the following reasons for asserting that short-term psychotherapy would receive increasing attention from therapists and researchers:

(i) Most forms of psychotherapy, whether or not they are specifically designated as short-term are, in fact, time-limited. For instance, it has been found that in clinical work, because of practical considerations, the average length of therapy is only a few sessions (Garfield 1978).

(ii) The evidence suggests that time-limited psychotherapy is as good as unlimited or long-term therapy (Luborsky, Singer, and Luborsky 1975).

(iii) In terms of patients' expectations, resources, motivations, and practical considerations, it is essential to develop psychotherapies that yield significant returns in the shortest possible time with the least expense.

Garfield (1977), too, has pointed out that short-term therapy should be the treatment of choice for practically all patients. On the basis of many reports (e.g. Harris, Kalis, and Freeman 1963) it would seem that about two-thirds of patients respond positively to such intervention; hence, if indicated, the remaining one-third can continue to receive therapy, can be referred elsewhere or, if the patients are considered to be unsuitable for the type of therapy currently available, treatment can be discontinued. One can compare the same type of treatments of different duration, or one can compare different

types of treatment of different duration. Miller and colleagues (1972) compared different treatments of the same duration in children and found that the treatments gave similar results.

FREQUENCY OF CONTACT

There is recent evidence that more frequent sessions of psychotherapy are more effective than less intensive therapy. Heinicke and Strassman (1975) compared psychoanalytic psychotherapy administered to nine-year-olds once a week, with similar therapy given to them four times a week. They reported impressive results for the latter on one- and two-year follow-ups when studying clinical data and reading ability. However, the samples were of small sizes and the evaluations appeared not to be blind. There were only few comments on the nature of the conditions being treated, their severity and duration, and no information was given about the characteristics of the therapists used in this study.

TIMING OF FOLLOW-UPS

Until recently this topic attracted only minimal interest, but it is, in fact, of the utmost importance that follow-up studies should be carried out at the right time if subtle therapy-induced changes in behaviour are to be detected. Recent advocates of detailed single-case studies using time series analysis have emphasized the importance of repeated measures over time in understanding the process as well as the outcome of therapy.

An interesting example of the importance of timing in follow-ups is given in the review, by Wright, Moelis, and Pollack (1976), of previous studies of individual child psychotherapy: they reported a consistent trend for improvement to be more pronounced at follow-up than at termination of treatment.

The psychosocial environment

This is a component of psychotherapy that has been almost entirely ignored in the literature on adult psychotherapy but that is less easy to pass over with children. It is now abundantly clear that children's environments at home (Rutter 1971; Birch and Gussow 1972) and school (Reynolds, Jones, and St Leger 1976; Rutter *et al.* 1979) have a profound effect on their behaviour and development. Any specific intervention occurs against the background of this development and is likely to interact with it in numerous ways (see the section on spontaneous remission later in this chapter (p. 31)).

Control groups in evaluative research

Intimately related to the questions of timing and psychosocial environment is that of control groups. In any study the control group must be selected with the object of the study clearly in mind. Even so, it is not always possible to select the perfect control group. There have been many accounts of the advantages and disadvantages of various control groups (e.g. Campbell and Stanley 1966), and some of these are discussed briefly below.

NO-CONTACT CONTROL GROUPS

This type of control necessitates the collection of data on subjects without their having knowledge of the proceedings, so that there is minimal experimental bias involved (Mitchell and Ingham 1970). There has been some research to demonstrate that most other control groups, including waiting-list control groups, show significantly more improvement than the no-contact control group.

NO-TREATMENT CONTROL GROUPS

There is evidence (Bergin 1966, who cites other studies) that subjects who do not receive psychotherapy seek help elsewhere: from friends, clergy, relatives, and professionals other than psychotherapists. This sometimes occurs in over 50 per cent of 'untreated' patients (Saslow and Peters 1956). For this reason a review of spontaneous recovery rates (Bergin 1971) suggested that such rates may be considerably lower than the 67 per cent suggested by Eysenck (1952). Bergin estimated the average spontaneous recovery rate for neurosis at 22.4 per cent, and he also pointed out that the rates reported by Levitt may have been inflated by including slightly improved cases.

Perhaps the most important work is that of Lambert (1976), who has demonstrated that testing sessions, or even one interview, may be of therapeutic benefit for the client: thus, the clients in no-treatment control groups may, in fact, be receiving some 'treatment'. Indeed, in some studies there were clients who attributed their improvement to the initial interview. Other studies have also shown that initial testing and initial contact have positive effects on clients, but, unfortunately, we can have no idea of how much help was being sought, nor of the extent to which these control subjects changed their behaviour in relation to the help that was being given. We can only surmise that some amount of treatment may be more effective than none, but, at the same time, it is difficult to visualize a situation where there really is no treatment being received at all.

WAITING-LIST CONTROL GROUPS

There are many advantages in using such clients as they are a common-place aspect of many clinics and thus form a natural control group; they are also motivated to complete the post-assessment programme by the guarantee of subsequent therapy. Unfortunately, though, like the 'no-treatment' control group, they may be seeking treatment elsewhere.

ATTENTION-PLACEBO CONTROL GROUPS

These groups comprise individuals who regularly meet the therapist for a chat, or for play, but who do not receive any proper treatment. The group may serve to control such factors as frequency of contacts, expectations of improvement, and therapeutic interest. In one study (Paul 1966) the attention-placebo group improved more than the no-treatment controls and, at the end of a two-year follow-up, this pattern re-occurred.

OTHER TREATMENT GROUPS

The groups that drop out from treatment are sometimes called 'terminator controls'; they are inadequate as controls because there is some unknown selection factor operating in termination. However, they can teach us some useful things about the therapy programme. It has been shown (Shapiro and Budman 1973) that clients from individual and family therapy who terminated against professional advice disliked the therapists because the latter were inactive, detached, and uninvolved, and the therapy did not have a clear direction. Those who remained in treatment tended to appreciate an active therapist who presented a programme with clearly articulated goals and direction.

CONTAMINATION OF CONTROLS

The problem of contamination arises when controls have contact with treated cases, as such contact allows for the possible transmission of treatment's beneficial effects. A similar situation occurs where contamination is transmitted through teachers or other workers, a matter of the greatest importance to the present study (see Chapter 10).

SPONTANEOUS REMISSION

So-called 'spontaneous improvement' is mentioned in many reviews of treatment outcome (e.g. Levitt 1971). This is a sensible recognition of the fact that just because improvement in a disorder follows treatment it does not necessarily mean that the treatment was

effective – a control group is necessary against which to compare the extent of improvement.

It is most important to recognize the statistical phenomenon of regression. If a group of individuals is selected by an extreme score on any assessment procedure that group will inevitably give a less extreme score if this procedure is repeated. In our research adjustments were made to allow for regression effects; these are more fully described in Chapters 3 and 9 and Appendices 2 and 3.

Second, many factors, apart from treatment, have been shown to be predictors of a good or bad outcome. For example, Richman (1977), in a one-year follow-up study of disturbed three-year-old children, found that a problem was more likely to persist if it was initially severe, if the child's parents had a poor marital relationship, if his or her mother was depressed, or the child had been subjected to poor housing conditions or stress during the previous year. These findings suggested that there are many mechanisms involved in whether a disorder improves or not: indeed, it may be that the treatment itself is a relatively weak change agent compared with the child's other social circumstances.

A third point to consider is that, in the light of recent research findings, it seems possible that professional psychotherapists do not have a monopoly of 'natural' therapy skill. Families do not only, or even usually, come to professionals for advice. Instead, friends or relatives with the right personality attributes (Truax and Carkhuff 1967) may effectively give counsel and advice, thus adding their contribution to the child's 'spontaneous' improvement (Bergin 1971).