Solution form therapy – results of NLP-based psychotherapy

(Überarbeitete englische Fassung der Tampere Studie)

Im Original (finnisch) wurde die Studie 2004 publiziert in:

Ojanen, M.; Kotakorpi, S.; Kumpula, S.; Tankklu, M.; Vadén, T.; Vikeväinen-Tervonen, L.; Kiviaho, M. & Hiltunen, S. (2004): *Terapiastako ratkaisu – NLP -perustaisen psykoterapian tuloksellisuudesta*. Helsinki: Mielikirjat.

About the authors:
Markku Ojanen is professor emeritus in psychology.
Martti Tenkku and Leena Vikeväinen-Tervonen have PhD in psychology. Both are trained NLP-therapists.
Sinikka Kumpula and Totte Vadén have M.A. degrees in psychology and are NLP trained therapists.
Soile Kotakorpi has Licentiate degree in psychology and is NLP trained therapist.
Matti Kiviaho has Licentiate degree in psychology, was Lecturer in the University of Tampere
Sinikka Hiltunen, is a doctoral candidate in cognitive science, and trained NLP therapist.
Tenkku, Kotakorpi, Kumpula and Vadén treated 54 clients.

The results of Tampere NLP-study
Follow-up of 62 clients participating in Neuro-Lingvistic Programming -based psychotherapy

Introduction

The project was originated at 1996. Its aim was to evaluate the results of psychotherapy based on Neuro-Lingvistic Programming (NLP). The applications of NLP have steadily increased in Finland during the past 15 years. NLP is still a controversial therapy in the sense that the members of old therapeutic schools do not accept it as an effective and valid psychotherapy method. Its image is somewhat esoteric and mysterious. Some critics include it within the large group of alternative treatments. Clearly the major problem is lack of empirical studies of its effectiveness. The present literature on NLP centers on theoretical analyses and case studies. We are not familiar about studies that fulfill even minimal conditions, such as: (1) adequate sample size, (2) valid and reliable measures and (3) at least two measure points (pre-post).

A reader who is not familiar with NLP is referred to the following publications: Dilts (1990); O'Connor & Seymour (1993). It is not easy to classify NLP within existing psychotherapeutic schools. Perhaps it comes nearest to experiential psychotherapies. NLP therapists talk a lot about beliefs which reminds of cognitive or solution focused therapies. There is a high emphasis on language; we live in constructed world. According to O'Connor & Seymour (1993, 1) "NLP is the art and science of personal excellence." If somebody has done something in a good way, it can be modeled, say these authors. Proponents of NLP want it to be something more than just a method of psychotherapy. It is a way to get things better anywhere in life.

This may be too good to be true, but it is easy to pick similar promises from older schools of psychotherapy. These kinds of promises are a minor problem, but much more serious problem is nearly total lack of studies related to NLP, particularly its effectiveness and efficacy. Without this kind of scientific support therapy is delegated to a questionable group of alternative treatments.

The origin of the study became by the NLP therapists working at the city of Tampere. Department of Psychology at University of Tampere had two clinical teachers who were invited to participate. The four major therapists in the study represented the following institutions: Tampere Health Center, The Family Federation of Finland, University affiliated Student Health Center and Tampere Psychological Center (private clinic). The first author was asked to design the study and select the research methods. He is a research member of local cognitive-behavioral therapy association, which means that he has no allegiance to NLP. The first author adheres to the contextual model described by Wampold (2001). Wampold shows persuasively that (1) clinical trials based on medical model have not been useful, (2) research should uncover the common or general effects, (3) emphasis on treatment manuals should be relaxed, and (4) focus should be on effectiveness rather than efficacy. "Effectiveness ... refers to the benefits of psychotherapy that occur in the practice context - that is, how effective are the treatments administered to clients who present to therapists in the community?" (Wampold, 2001, 61).

The precondition for the study was that it should harm the therapeutic process as little as possible. Therapy was always master and research its servant. It was not possible to design a research therapy -based study (see Matt and Navarro, 1997, 24). With this precondition the control group could not be included and the client group became very heterogeneous. It was not possible for ethical reasons to have a part of the clients wait, because the condition of many if not most of the clients was quite acute. The major asset of the study was that it is true to actual clinical practice. The basic aim of study was to evaluate comprehensively the results of NLP-based therapy. Both immediate and long-term follow-ups were included in the design.

The apologies of within-group designs can go to far, as Hunter and Schmidt (1990) have persuasively written. The within-group design has much more statistical power than the between-group design and its main effects (change scores) can be unambiguously interpreted. They even state that "under most conditions, the within-subjects design is far superior to the between-subjects design" (p. 339). Important data is lacking in the latter design, because gain scores are missing.

There clearly are problem behaviors for which within-group designs are suitable. When a problem has been shown to be stable for about the same period as the follow-up period is, the within-group design is a good alternative. On the other hand, if a problem is highly unstable and corrects itself in a very short time, within-group design is not good. In reality the therapy clients usually do have quite stable problems, because they do not immediately seek help. In most cases immediate help is not even available.

The major reason for lambasting within-group design in psychotherapy is the plausible expectation that clients in therapy groups "spontaneously" get better. They seek treatment when their condition is at worst and if they do not get the therapy researchers offer, they may seek something else. The Lipsey-Wilson (1993) study shows that control/comparison designs give lower effect sizes (ES) than within group (pre-post) designs. Thus one must subtract something from the effects of within-group results. The big question is: How much? The mean ES for between-group designs was 62% of mean within-group design ES. Can we base our estimate for spontaneous effects for this percentage?

The mean ES of Lipson and Wilson's pre-post designs was 0.76 which included all kinds of psychological and educational effects. The ES for control/comparison studies was 0.47. The spontaneous recovery -effect would thus be 0.29 or about 0.30. As a percentage spontaneous recovery is 38%. However, the effects of psychotherapy are considerably higher, about 0.80 -0.85 (Wampold, 2001). Is spontaneous recovery of 38% a viable estimation in psychotherapy? This would mean ES of 1.29-1.37 for within-group psychotherapy studies. It is probable that problem groups improve spontaneously, but how large this effect really is? How much spontaneous recovery depends on the quality of presented problems?

Matt and Navarro (1997) similarly stated in their review that using pre-post changes are expected to inflate ESs. The probable reasons are spontaneous remission, maturation, testing effects, statistical regression effects and attrition of the more severely ill. They mention 8 studies where pre-post ESs were compared to control groups. Unfortunately five studies concern juvenile delinquents which cannot be included among usual therapeutic groups. This leaves two studies where pre-post changes compared to control-therapy differences were larger and one where the result was an opposite one.

Greenberg, Elliot and Lietaer (1994) summarized the studies on eperiential psychotherapies. They calculated pre-post change ESs for 37 studies involving 1,272 clients. ES was 1.51 for early (1-8 months) and 1.26 for late (9-144 months) follow-ups. For controlled studies (16) SE was 1.29. The difference between pre-post and controlled studies was thus small. This group of studies seems to point to very low, almost nonexistent, spontaneous recovery.

Taylor (1996) conducted a meta-analysis for social phobia. In this study pre- to post-treatment d-statistic was computed for waiting list control, too. For 6 studies the mean ES was -0.127. Thus social phobias of waiting list controls did not change better. Here is an example of quite stable problems which often have lasted for years.

McDermut, Miller and Brown (2001) conducted a meta-analysis of group therapy for depression. In most studies untreated control groups improved somewhat. Their table 4 shows that treated clients changed their BDI values from 23.90 to 12.30 and untreated controls from 24.90 to 20.90. The amount of change in control groups is 34.5% of group therapy change. The average ES for 48 studies was 1.026. An estimated ES for within-group design would be 1.57.

There are studies that show the course of untreated mental problems. Posternak and Miller (2001) have collected results on wait-list control groups from depression research. The groups included had to have a diagnosis of major depression or high level of depression on BDI. The follow-ups were 2-20 weeks. The measure was either Beck Depression Inventory or Hamilton Rating Scale. The final meta-analysis included 19 studies and 221 persons. The mean decrease in BDI was 15.7% and 11.9% in Ham-D. In BDI this meant a decrease from 24.3 to 20.5. The amount of decrease did not depend on the severity of the depression. 11 (18.3%) had achieved a final BDI scores less than 10 which can be regarded as normal. These changes would give ES of about .50 (if sd of BDI is between 7.0 - 8.0).

Nietzel, Russell, Hemmings and Gretter (1987) did a meta-analysis of studies concerning psychotherapy for unipolar depression. They used 31 studies with 1040 clients having BDI as an outcome criterion. 60 separate ESs were reported. In this study ES was calculated by using normal groups as the comparison group. The control group change in Z-scores was 29% of the treatment group change from pre-test to post-test. It can be calculated that treatment groups were before therapy at 21.36 on BDI and improved to 14.14 at post-test. The corresponding values for control were 26.33 and 22.27, which is very similar to results of Posternak and Miller (2001).

Borcovec and Ruscio (2001) analyzed psychotherapy studies on generalized anxiety disorder. The pre-post ESs for anxiety were 2.48 for cognitive-behavioral therapies, 1.72 for behavioral or cognitive therapies and amazingly 2.09 for placebo or alternative therapy. This must be a very mixed bag of studies; it is not very helpful combination. Waiting-list groups did not change at all. ESs of depression measures were about half of these.

In summary, the results about spontaneous recovery are still mixed ones. It is a real phenomenon, but depend on the nature of problems. We will return to this question later.

Effects of placebo are estimated to be 0.20 - .47 (Barker, Funk & Houston, 1988, 588; Lambert & Bergin, 1994, 151; Stevens, Hynan & Allen, 2000, 282; Wampold, 2001, 205). Wampold's value of .40 is accepted here as the best estimate. Lambert and Bergin (1994) estimated the ES for different components as follows:

psychotherapy versus no-treatment	0.82
psychotherapy versus placebo	0.48
placebo versus no-treatment	0.42

The major question concerns about the relative effects of common and specific factors. Meta-analyses based on medical model using clinical trials emphasize the role of specific factors. This component is sometimes estimated to be much stronger than common or general component, about 0.40 - .50 in SE units (Stevens, Hynan, & Allen, 2000), but the computations of Wampold (2001) which show that effects of psychotherapy are not specific but common to all therapies are accepted here as the most plausible. A major part of the effects are due to therapist, client and interaction variables beside common factors. It is extremely difficult to show the effects of specific factors, because hundreds of studies have not succeeded in this. An exception may be

very narrow or specific problems, like spider phobia which Öst (1989) has shown to have 100% recovery rate.

The results of therapy depend both on the nature of variables and problems. Generally more specific variables are easier to change especially if therapy has been targeted to them. Thus behavioral criteria can change dramatically, but quality of life variables change much less. If meeting new people is the aim of therapy, increase from 2 to 6 contacts per day is dramatic (200%) change, but in Quality of Life (QOL) or Subjective Well-Being (SWB) units change of 15-20% is very good. Specific, concrete symptoms can also change markedly. Phobias can be treated very successfully, but amount of changes in general anxiety or depression are often very similar to QOL or SWB variables. Some measures are general but change easily (mood), though changes may not last.

Measures are usually related to the quality of client problems. Specific fears and phobias are an example of narrow problems and the specific problem-related measures used by the therapists reflect this context. In phobias QOL measures should also be used. The level of QOL of such persons may not be on an average very low which means that changes in QOL are small. Depression and general anxiety are an example of comprehensive problems. In these problems specific symptoms are often measured (e.g. insomnia). Changes are very variable, because not all have insomnia and this may not be a special target of the therapy. Finally, comprehensive problems can and should be evaluated by global or general measures (SWB, QOL).

The problems of measure selection make the evaluations of psychotherapy results very difficult. The three domains (well-being, symptom, life function) used by Stevens, Hynan and Allen (2000) is a step into a right direction, but not entirely satisfactory. This classification mixes globality and stability of measures. The measures should be differentiated as follows (table 1)

Table 1. The classification of measures according to their globality and stability.

Globality	Stability: High	Stability: Low
High	QOL, SWB Satisfaction Traits A	Mood State measures B
Low	Specific fears Skills Habits C	Transient negative emotions and behaviors D

Even this is not a completely satisfactory classification. Some stable, but narrow habits, like smoking, are quite resistant to change, but specific fears are easier to change. It is not sensible to compare the ES's of smoking, snake phobia and depression in order to show the efficacy of a particular therapy. However, we can make some crude predictions about spontaneous change, therapeutic change and durability of therapeutic change based on table 1 as follows:

	Spontaneous change	Therapeutic change	Durability of change
A	Small	Moderate	Good
B	Moderate	Large	Quite good
C	Very small	Small - Large	Very poor - Very good
D	Large	Very large	Poor - variable

Of course, the change does not depend solely on measure, but always on problem, too. There must be of well-being deficit in the beginning of the study, because without it there can be no SWB changes at all. State measures are often used in depression and anxiety problems. Rarely the measures are pure state measures but purport to measure trait-like features of the problem. Thus the measures are A-B mixtures. Fear-related measures are typical in phobias. Therapeutic changes are often large and durability is very good. Skills can be narrow, but often training concerns a larger life function area like assertiveness or social skills. The globality of these is quite high, between A and C. Changes are usually moderate and stable. Finally personal habits, like eating, exercising and smoking, change very little spontaneously. Many habits are resistant to change and changes are not durable. Transient negative emotions and behaviors are usually known to be such and not in need in any therapy. If help is sought, change is large, but does not necessarily last.

In summary, the results depend on the nature of client problems and measures. The changes in complex or comprehensive problems are probably much difficult to obtain than in specific, less severe problems. Changes in global life quality measures are smaller than in specific problem-related measures. Two meaningful alternatives are available. (1) Study a particular problem group and use both specific and global measures. Comparison with other kinds of problems is not useful. This also means that general meta-analyses that combine various problem groups are not very meaningful. (2) Collect a large representative sample of clients having the most common problems. Use a wide variation of outcome measures. Very specific measures are not useful, because the variation of symptoms and problems is great. Global measures may underestimate the results in those cases where the problem was specific and therapy was effective.

Clinical significance. This concept refers according to Kazdin (2001, 455) "to the practical value or importance of the effect of an intervention - that is, whether it makes any real difference to the clients or to others with whom they interact in everyday life". The major criteria or measures for evaluating clinical significance of change are as follows (Kazdin, 456):
1. Similarity to normative samples at the end of treatment.
2. Statistical departure on a measure from score of a dysfunctional sample.
3. Amount of change from pre to post.
4. No longer meeting criteria for a psychiatric diagnosis.
5. Complete elimination of the problem or symptom.
6. Ratings of current functioning.
7. Whether the original problem continues to be evident or to affect functioning.
8. Whether the change or changes produced in treatment make a difference.
9. Change reflected on such measures as arrest, truancy, days missed from work, hospitalization, survival, and cost.

Because of the pre-post design of the study, these criteria are used to a great extent to evaluate the results.

The study

Therapists

The four major therapists (psychologists) were all very experienced having done therapy over 20 years. They all had Master level in NLP and had additional courses in many areas related to therapy (hypnosis, rehabilitation). In addition to these four other experienced therapists provided **8** cases, but could not provide more clients for the study. Although all therapists have a strong NLP allegiance, they are eclectic and use other methods, like hypnosis, which are not tightly associated with the core methods of NLP. The study was accepted by local ethical committee and the clients signed a form where they promised to participate in the study.

Clients

The aim was to include in the study all consecutive clients coming to the four therapists fulfilling the inclusion criteria. Those having clearly somatic reasons were excluded, and similarly those who during the first session were transferred to couples' therapy or to long term analytic therapy. These transfers were generally made according to the wishes of the clients. Those coming for only a consultation or counseling (one session) were also excluded because it was not possible to give them two questionnaires during the same session. Finally, those not wanting to participate were of course excluded. It can be estimated that about 40% of all new clients of the major therapists participated in the study during the study period.

The final study group included 62 persons of which 50 were women and 12 men. The level of schooling was quite high: 42 (68%) had completed high school (12 yrs of schooling). 22 (42%) either had M.A. level or studied towards it. 18 (29%) were married, 14 (23%) were cohabitating, 8 (13%) were divorced, 2 (3%) widowed and 20 (32%) were not married. Age varied from 18 to 54. Most of the clients were between 25-40 years. 34 (56%) had received psychological or medical treatment for their problem before the present therapy. 13 (21%) were on medication during the therapy.

The therapy process did not include a formal diagnosis, but each therapist made provisional diagnosis based on ICD 8 classification. Only those diagnoses were listed which the therapists felt were reasonable enough. The attitude of NLP therapists toward diagnosis is negative. It is not considered to be helpful for the therapy process. Diagnosis was given to 54 clients as follows (table 2). Sometimes more than one diagnosis was given. This means that the number of all diagnoses is larger than 54.

Table 2. The provisional ICD 8 diagnoses given to 54 clients.

	Primary	All
F32.0-1	9	9
F40.1-9	14	21
F41.0-2	11	12
F42.0	-	1
F43.21-25	8	8
F45.34	-	1
F48.8	1	1
F50.2-3	2	3
F51.0	1	1
F52.3-10	1	3

F60.0-6	3	4
F62.0-1	1	2
F.66.21	-	1
F95.1	-	1
G40.3	-	1
T74.2	-	1
Z61.0-8	-	2
Z63.0-4	2	3
Z73.0-3	1	3

In order to give a comprehensive view of the problems the client answers to two questions "What are those things, factors or happenings which have brought you to therapy or discussions" and "What is the most important or central problem, which has brought you to therapy or discussions". The stated problems are given in shortened version in table 3. It can be seen that the problems really were manifold. Some appear not to be severe, but some were very serious.

Table 3. The problems that brought the clients to NLP-based psychotherapy.

- death of spouse
- work exhaustion
- anxiety, loneliness
- tight nerves, physical symptoms, negative attitude to life
- depression
- difficult childhood memories
- poor self-esteem, eating problems, negative attitude
- marital violence and marital problems
- sexual harassment and fears related to it
- depression, weakness
- insomnia, bad dreams, sexual violence, incest
- anxiety and somatic symptoms, incest
- tic symptoms, social fears
- depression, difficult childhood memories
- fear of desertion
- social fears and anxieties
- panic attacks and somatic symptoms
- lack of strength in studying, models coming from childhood are not working
- panic attacks, social phobias
- jealousy
- insomnia and stress, exhaustion
- public speaking anxiety, social anxiety, sexual problems
- depression, insomnia, exhaustion, self-critique
- somatic symptoms, fears and bad feelings
- fear of dental operations
- exhaustion, fears
- problems of self, attacks of fear, sexual problems and problems with father
- public speaking and social anxiety
- writer's cramp with related psychological symptoms
- anxiety, social problems, fear of failure, lack of worth
- attacks of anxiety especially during nights
- starting studies, anxiety

- social anxiety, depression, dependence on father
- fears related to getting work, unexplained anxiety attacks
- study problems, problems in concentration and decision making
- red cheeks and feelings of guilt
- social and study anxieties
- study problems
- depression, panic attacks
- passing out during teaching
- insecurity about future and study subject, low self-esteem
- wish of parents and hopelessness, eating problems
- conflicts with father and boyfriend, bitterness toward father
- problems with studying and professional identity
- depression, which is getting better, anxiety
- low self-esteem, jealousy
- anxiety and panic are back, fear about future work
- difficult relationships with spouse and relatives, lack of independence
- bulimic symptoms and low self-esteem, study problems
- lack of orgasm
- work exhaustion
- co-dependence, fear of failure, depression, anxiety
- anxiety, too much energy, alcohol problems
- nervousness, work fears, things getting worse, panic attacks
- difficult childhood experiences (violence), asthma, illness of children
- problems are between ears - may get physically ill, violent childhood memories
- mental health, marital problems, money problems
- ill feeling, nothing interests, alcohol problems, marital violence
- fear attacks at nights

The presented problems at the first meeting were categorized as follows:
Specifity-globality
I Specific, narrow symptom or problem (fear of dental operations)
II Neither specific or comprehensive, but clearly has an effect on a certain life area (social anxiety, study problem)
III Global or comprehensive symptom or problem (depression, lack of self-esteem, exhaustion)
Symptoms-problems
I Only symptoms mentioned (anxiety, depression, pain)
II Both symptoms and problems
III Only problems mentioned (problems with a specific person, life event or change)

These classifications were done together by the first author and an independent evaluator.

The results of this classification were as follows:
Specifity-globality

I	2
II	22
III	38

Symptoms-problems

I	5
II	45
III	5

There was only one client who presented a specific symptom (fear of dental operations) and even this may not actually be a narrow problem. The majority of the clients described their problems in general or global ways. In a second classification both symptoms and problems were usually given, like depression and problems with social relationships. Of these classifications the first one (specifity) shows statistically significant differences between the two large groups (II and III) both on SWB and problem severity. Those who use global expressions have lower SWB and more difficult problems.

The client situation can also be described by comparing their ratings of Subjective Well-Being (SWB) measures described later. The results of 6 show that the level of well-being of the clients is markedly lower than that of general population.

The measure often used by NLP therapists requires the clients to use a scale having the following end-points: 0=the worst possible situation you can imagine and 100=the best possible situation you can imagine. What is your situation now? The mean value given by the clients was 45.5 (sd 18.9).

The average number of sessions was 8.2 with a range of 2 to 23. Number of sessions had quite low correlations with the client condition. For instance, the client total SWB (subjective well-being) correlated -.06 and total PE (difficulty of presented problems rated by therapists) -.22 with number of sessions. The therapist rated well-being (TWB) correlated .33 with number of sessions.

The design of the study

The design was of one-group pre-post or within group -type. The measuring points and number of respondents (in parenthesis) were as follows:

I The before therapy measures (62)
II The after therapy measures (61-62)
III Follow-up 6 months (51)
IV Follow-up 1,5 - 2 yrs (41-42)

The therapeutic changes could be compared with changes in five groups. Within these groups one can see what happens to persons, whose level of SWB is similar to the present client group. All groups had two measuring points, but the time differences varied. These groups were
A 222 employees from a large Finnish electronics company. Originally a random sample. A follow-up of 1,5 years. Percentage of women was 28.
B. 144 university students (University of Tampere), a random sample. A follow-up of 4 years. Percentage of women was 60.
C. 107 personality psychology students at University of Tampere, 4 month time difference.
D. 114 personality psychology students at University of Tampere, 2 month time difference.
E. 66 personality psychology students at University of Tampere, 4 week time difference. The last three groups were from lectures given by the first author 1999-2001. The students had majors in many different subjects. The percentage of women at these courses was about 90.

Beside these a group consisted of 45 unemployed, MA-level trained persons of which 21 participated in 2.5 - month counseling sessions in order to improve their chances in getting work (Vikeväinen-Tervonen, 1999). The counseling did not statistically significantly improve SWB

and thus the whole group can be used here. Of the participants 41 (80%) were women and the age-range was from 25 to 55.

From all these groups those persons could be selected whose SWB values were under limits known to be typical of problem groups having mental problems.

In addition a comparison group of 238 of persons can be used. The mean age of this group was 39.7 yrs and it had equal number of men and women. This is a convenience sample, but is used here because all of the scales shown below were utilized in this group. Its results were highly similar compared to random samples of university students, employees of a large Finnish company and employees of a hospital district.

Measures

The measures were primarily based on the clients' own ratings of their psychosocial well-being and problem cognitions. Therapist ratings were also used, though it was recognized that these were the most reactive measures available. Thirdly, the clients rated the usefulness of therapy. And fourthly, follow-up data about mental health service use and medication was collected.

The following types of measures were used:

	Number of items	Scale range
I Clients		
1. Descriptive visual analogue scales	13	0-100
2. Problem experience ratings; adjectives	25	0-10
3. Ratings of change after therapy	11	0-10
4. Therapeutic relationship, satisfaction and fulfillment of aims (after therapy)	3	0-10, 0-100
II. Therapist		
1. Descriptive visual analogue scales	6	0-100
2. Problem quality ratings and expectations (first meeting)	22	0-10
3. Ratings of change after therapy	11	0-10
4. Therapeutic relationship, satisfaction and fulfillment of aims (after therapy)	3	0-10, 0-100
III Other follow up-data		
1. Unemployment	yes/no	
2. Sought help	yes/no	
3. Medication	yes/no	

The major precondition for selection of measures was usefulness from the point clinical work. This meant that the measures had to be comprehensive and practical. Subjective well-being was measured by 13 descriptive visual analogue scales (DVAS) devised by senior author (Ojanen, Tuori, & Lauren, 1997; Ojanen, 2000; 2001). There are about 40 SWB or QOL scales available at present. They have been used in numerous studies in Finland since 1980 (Ojanen, 2000). The mean levels and distributions of both normal and many problem populations are known. All DVAS scales have 0-100 range. 100 denotes a very high level of a certain attribute and 0 a very low level or lack or something. A few similar scales are in use. In VAS-scales 0-100 range is also

used, but only the endpoints are described. Global Assessment Scale (GAS) in DSM-classification is alike to the DVAS scales. Sperry, Brill, Howard and Grissom's (1996, 83-4) Life Functioning Scales (LFS) appear to be much the same as present scales. The six LFS scales have also a range of 0-100. The internal consistencies and reliabilities were also similar to the present ones. An example of these scales is given in appendix 1.

SWB and QOL measures are in practice often very similar. In our opinion QOL concept is more comprehensive and should include variables related to economic situation and environmental quality. The scales used here concern more about personal well-being.

The 4 week stability of the individual scales has varied between .66 and .82 and the stability for the sum of 5-7 scales has been . 88 - .92. (Ojanen, 2002). Both client and therapist used these scales. As in previous studies (see Ojanen, 2002) the 13 DVAS scales load quite highly on first unrotated principal axe. The 7 scales used here for total SWB are well-being, life satisfaction, working ability, self-confidence, life control, anxiety and mood. The names of other scales used can be seen from table 6. In the present study Cronbach's Alpha for total SWB was .81 at the first and .87 at the fourth measurement. The retest correlations of individual scales in the present study were quite low because of the therapeutic intervention. The average scale correlation between before and after measurements was .57. When the after therapy measures were correlated with 6 months follow-up, the average scale correlation was .53.

For the therapist rated well-being (TWB) six DVAS scales, anxiety, aggression, mood, vigor, relaxation and working ability, were used. The first five loaded highly on the first unrotated factor. The sixth scale, working ability, had lower loading than the other scales. The reliability (Cronbach's Alpha) of the sum scale (including working ability) was .87 at first and .85 at second measurement.

Problem experience ratings were devised on the basis of project group discussions. The scales aimed to measure those aspects that are important from the point of NLP. Of these 9 were selected for the final follow-up (Cronbach's Alphas were at 1st measurement .66 and at 4th .67). Problem ratings were not onedimensional; the factor structure was not very clear. These 9 items loaded mainly on two factors, which were correlated: worsening-disturbing and helpless-hopeless. The ratings of therapist covered two areas: good contact + optimism (7 items) and problem severity (4 items). The rest of the items described the relationship from the point of NLP. The results of these variables are not reported here.

Client and therapist ratings of change after therapy had 11 alternatives from 0=very negative change... through 5=no change ... to 10=very positive change. The areas rated were as follows:

1. Basic problem
2. Self-confidence
3. Functional ability
4. Relationships
5. Stress taking
6. Satisfaction
7. Growth
8. Mood
9. Anxiety
10. Life control
11. Total life situation

Clients rated these areas three times (II, III and IV) and the therapist once (II). The Alphas for total change were .95 for clients (II) and .94 for therapists (II).

Hypotheses

It can be expected that NLP-based eclectic psychotherapy does produce positive results. Though the first author leans toward behavior-cognitive therapies, he favors the contextual model described by Wampold (2001). The therapists had a long experience and a high allegiance to NLP. Thus we expect that

1. Very good working alliance is formed and it is rated to be useful.
2. Mean level of SWB improves significantly.
3. Cognitive components tailored to NLP change more than SWB.
4. After therapy the level of SWB of the clients approaches the level of normal comparison groups.
5. The obtained results will last at follow-ups.
6. The changes are faster than in available comparison groups.
7. The effects of NLP -based therapy are of similar size than those of the other therapies.
8. The changes should be more than in those groups having as difficult problems but which do not receive therapy ("spontaneous change").

The strong points of NLP system are very similar to those of other therapies. The therapists have high allegiance to NLP, they use many social influence tactics and are able to create positive expectations. Positive expectations will be aroused by a large package of specific therapeutic methods or tactics. A list of 36 different NLP techniques was provided by Dr. Martti Tenkku. The client can thus expect that if the first method tried does not work, one of the 35 will! The NLP therapists are exceptionally eager to try any method regardless its origin; they are positively eclectic, which favors common factors. One thing is sure: their work can not be systematized in manuals. One therapist uses behavioral methods, another hypnosis and a third one leans to psychodynamic concepts.

In summary, in order to show that the results are positive, the outcome of the study should be better than what can be expected by spontaneous recovery.

Data analysis

The pre-post differences were analyzed by matched pair t-tests and comparisons with normal groups with independent group t-test. Emphasis is on scale sums, but results on individual items or scales will also be reported. One-way repeated analysis of variance was used to analyze follow-up measures.

There were a few instances of missing data, in which cases this variable was not included in statistical analyses.

Results

Status of clients before therapy

Well-being deficits of the clients. The before therapy ratings of the clients were markedly lower than those of the comparison group (table 6). However, the differences between client and comparison group were not very large in absolute sense. The clear majority of the clients were working or studying. Some were unemployed (7), but no-one was on full social security or pension. Thus the clients do manage, but have a lowered level of well-being. Other studies (Cummins, 1996; Ojanen, 2000) show that problems must be really dramatic before SWB was below the neutral scale point. Even hospitalized psychiatric patients usually have an average level over 50. Though scales are not comparable in any absolute sense, it can be seen that the first five scales cover the most problematic areas of life. Satisfaction with life and general mood were markedly lower than in comparison group. On the other hand their life was secure and their social life was quite good. The variation was large in the client group. There were those whose specific problems were not reflected on general well-being. The distributions of total SWB of the client and comparison groups are shown in figures 1a and 1b. There were quite many clients whose level of SWB was clearly on the normal range.

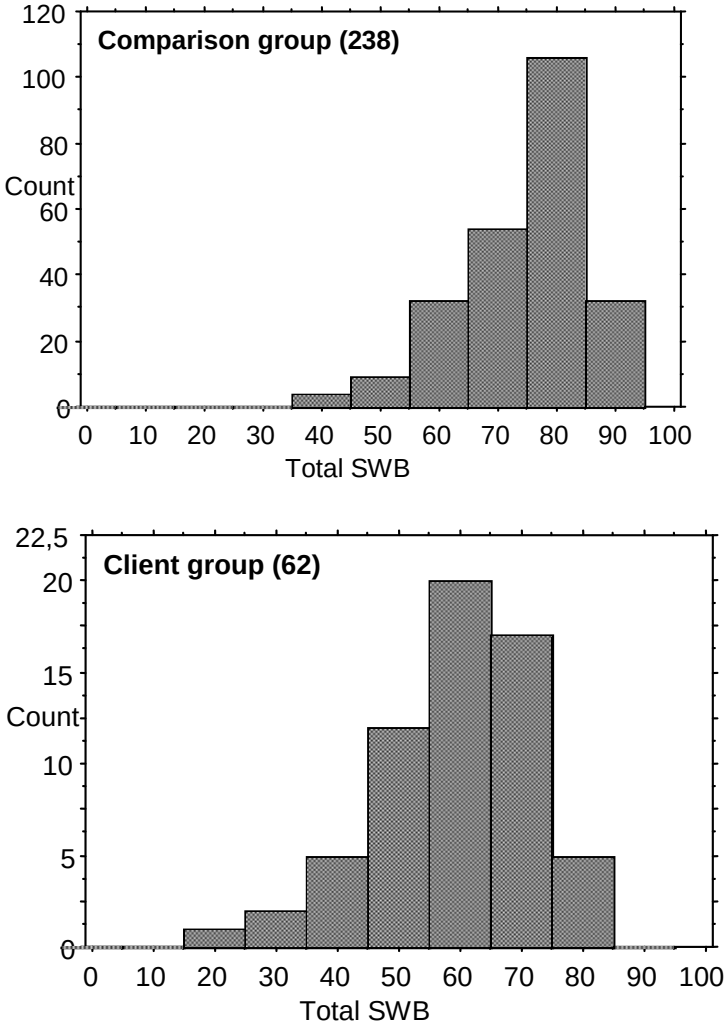


Figure 1. The distribution of total SWB variable in client group before therapy and in healthy comparison group before therapy.

Group differences in SWB. There were statistically significant differences ($p<.01$) between men and women in self-confidence, anxiety, life-control and physical condition. The means of men were more positive. The difference was between 5-8 scale points. There were no significant differences in total SWB. Those having spouse or partner had higher total SWB than those who did not ($p <.03$). Clients who had sought help before coming to the present therapy had somewhat lower level of total SWB ($p<.005$) and those using global descriptions about their problems had also lower level of total SWB ($p<.02$).

Problem experience. The results on nine scales used in all follow-ups are reported here (figure 2). All means except worsening were below the middle point of the scale, where the experience was of "both-and" -type (e.g. both disturbing and not disturbing). The problem was highly disturbing, practically always in mind and induces helplessness. It was not worsening, but rather quite stable (somewhere between worsening and not worsening). Figures 3a and 3b show the distributions of these two scales. It may be that just coming to therapy gives reason to optimism though the problem was disturbing.

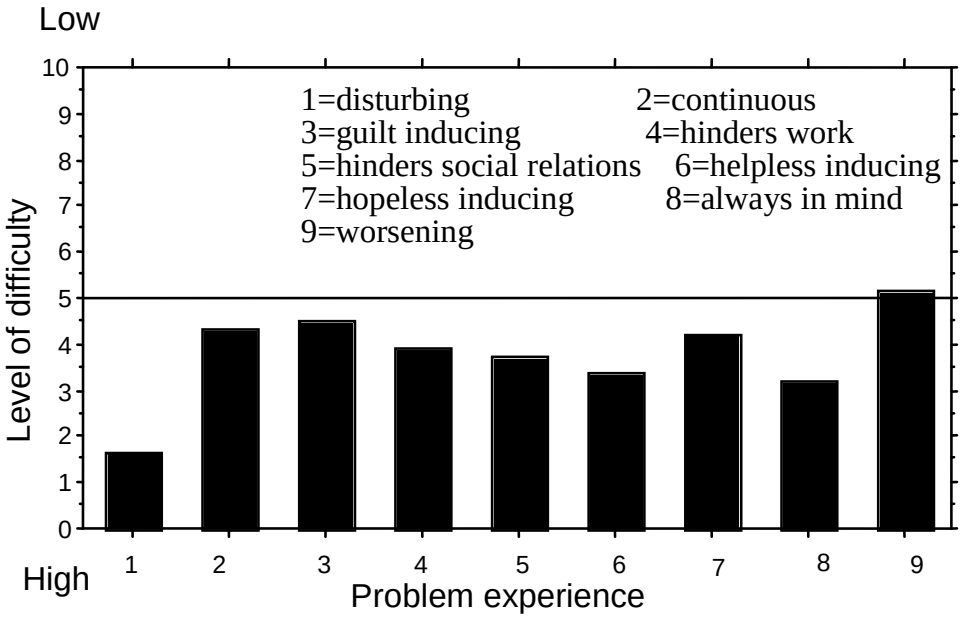


Figure 2. The pre-therapy means of nine problem experience scales.

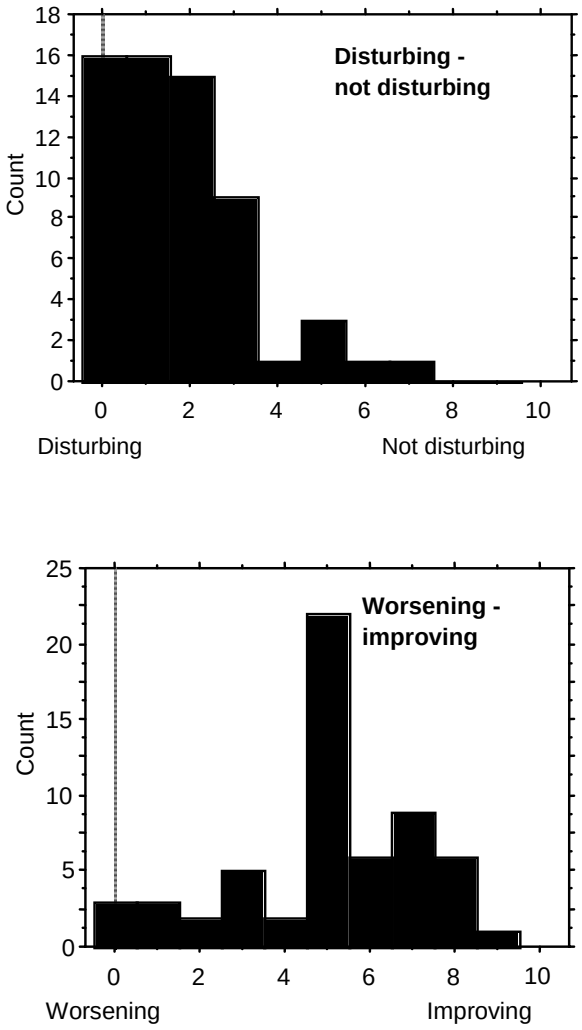


Figure 3. The distribution of two problem experience scales before therapy.

The therapist sees the problem as quite difficult (mean 3.6, sd 2.1) and not so easy to solve (mean 5.8, sd 2.0; see the scale direction), though the variance is again large (see figures 4a ja 4b).

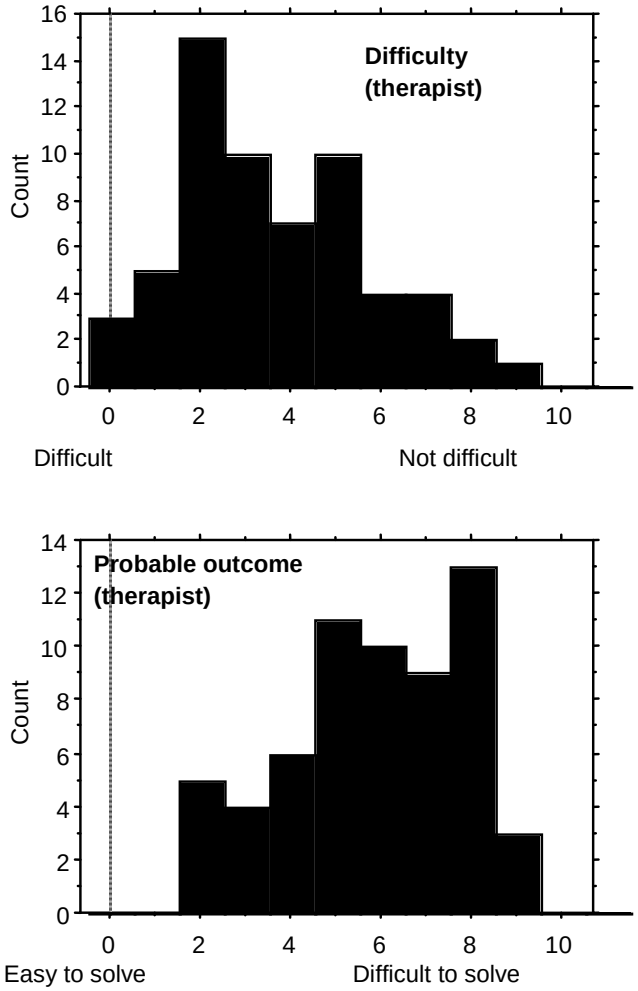


Figure 4. The distribution of two problem scales by therapists before therapy.

The situation of the clients can be summarized by using the 5 measures which describe the severity or difficulty of the problems. The 5 measures are
1. Severity of the problems rated by the therapists (one 0-10 -item)
2. Problem experience ratings by the clients (9 0-10 items)
3. Severity of the present life situation rated by the clients (one 0-100 -item)
4. Ratings of Well-Being by therapists (6 0-100 items)
5. Subjective Well-Being ratings (7 0-100 items)
0-10 scales were transformed to 0-100 scale.

Figure 5 shows that therapist severity ratings were the lowest ones which means they rated the client problems quite as severe. Client problem ratings were also quite negative. General well-being ratings were on the positive side. Of course, the scales are not comparable in the strict sense, but these means probably show that although problems can be fairly severe, but general well-being was still adequate.

If we choose 50 as a cut-point and count those clients who were at 50 or under it on a particular variable, the numbers are in the order of bars as follows: 54, 49, 36, 22 and 14. There were 4

clients who were over 50 on all variables. Their reasons for coming were something special and needs some comments. They have described their problems as follows:
- start of studying and
- co-dependence
- fainting during a lecture
- no clear reasons

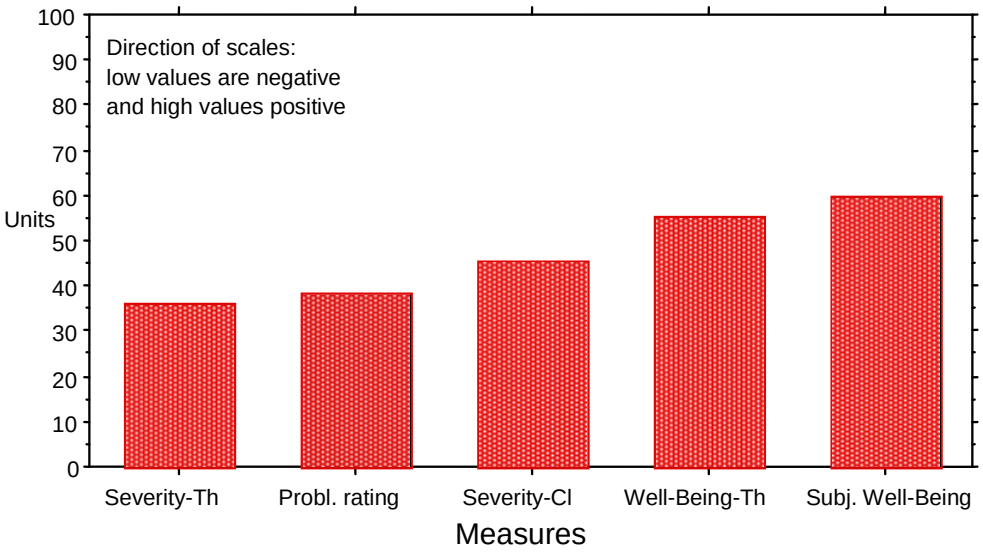


Figure 5. Severity of client problems and life situation before therapy

Comparison with other clinical groups. Muurinaho (2000) studied 189 persons who participated in self-help groups offered for persons having serious problems. Only 18% of them were able to work and 55% had been in mental hospitals. 62% were at present in some kind of treatment. The group was very heterogenous, but included depressive persons as a largest group. The percentage of women was 61 and the average age was 47 years. Table 5 shows means and standard deviations on those DVAS scales present in both studies. There were no significant differences on mood and self-confidence, but anxiety of NLP-group was higher than among self-help group participants. Lower working ability and weaker social relationships of the latter group reflect the chronicity of their problems.

Table 5. Comparison of NLP study (N=62) with self-help group participants (N=189) on Subjective Well-Being. Means, standard deviations and statistical differences on five DVAS scales.

		This study	Self-help group	t	
1.	Mood	56.0	53.0	0.998	
		19.7	22.4		
2.	Anxiety	53.2	46.0	-2.804	**
		15.7	21.9		
3.	Self-confidence	58.5	58.0	0.167	
		20.2	20.5		
4.	Working ability	72.0	48.0	9.046	***
		18.3	29.3		
5.	Social relationships	75.7	69.0	2.354	*
		18.4	21.9		

* p<.05, ** p<.01, *** p<.001

Experience of therapy

The experience of therapy was very positive. Figure 6 shows the means of three identical scales for both clients and therapists. The aims were fulfilled, the relationship between client and therapist was good and both parties were satisfied about the therapy. The ratings of the clients were somewhat more positive than those of the therapists. The correlations with client and therapist ratings were as follows in the above order: .67, .58 and .31. There seems to be quite a good agreement about aims and alliance, but not about satisfaction. It is of course possible and even probable that fulfillment of aims has been discussed after therapy. On the other hand relationship and satisfaction probably was not discussed and may even be an evaded topic in therapy.

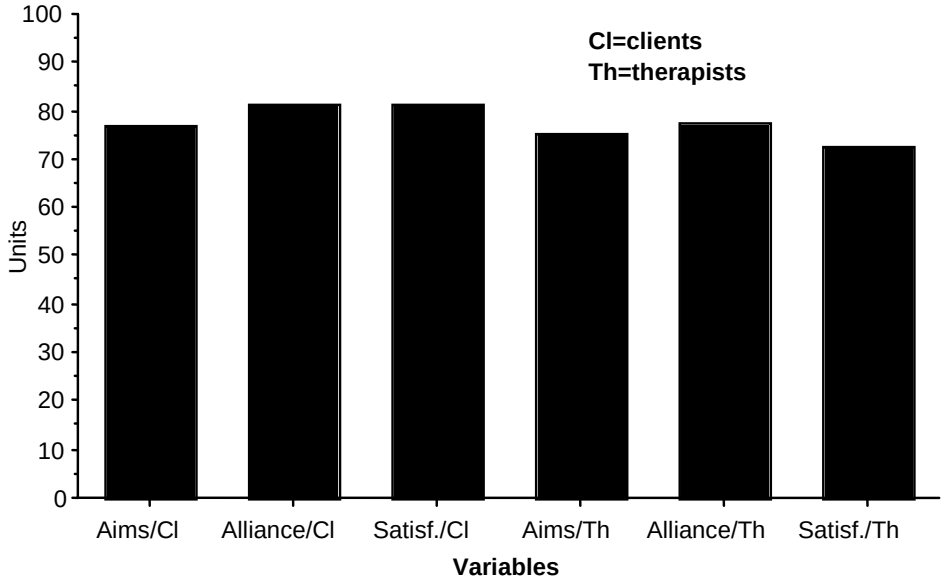


Figure 6. The mean ratings of clients and therapists on three therapy variables (fulfillment of aims, therapeutic relationship and satisfaction with therapy).

Effects of therapy

The major outcome variables were total SWB consisting of 7 DVAS scales, total TWB (6 DVAS scales) rated by the therapist and total problem experience (PE) by clients (9 items). The correlations of these were .62 (SWB and PE), .45 (SWB and TWB) and .31 (PE and PE). Changes on individual scales are also reported in the following tables.

Subjective Well-Being. The before and after means of the NLP -based psychotherapy group are shown in table 6. The table also includes the means of the comparison group. The changes in SWB variables were without exception statistically highly significant. The variables were, of course, highly correlated and thus statistical tests were not independent. However, the original levels of SWB variables were different and the amount of change was also different. The results of individual scales are thus meaningful to show. For example, the major problems were just on those areas that could be expected: mood, anxiety, well-being, self-confidence and life satisfaction. The changes in these were the greatest. On the other hand, physical condition, activities, social contacts and security approached the normal level before therapy. Marked changes were not expected in these.

Table 6. The means and standard deviations of DVAS scales before and after therapy of the client group (62) and the comparison group (238)

		Clients		t-value	Comparison	t-value
		Before	After	(change)	group	(differ.)
1.	Mood	56.0	68.8***	5.61	71.5***	6.96
		19.7	11.1		14.3	
2.	Anxiety	53.2	38.3***	7.43	32.2***	8.17
		15.7	11.1		17.6	
3.	General well-being	62.0	78.7***	6.81	78.4***	6.89
		19.2	12.3		15.9	
4.	Self-confidence	58.5	74.3***	8.38	73.9***	6.88
		20.2	13.5		14.4	
5.	Life satisfaction	55.1	73.2***	8.70	74.1***	9.71
		16.8	11.6		12.9	
6.	Optimism	60.6	67.2***	4.27	68.5**	3.52
		15.7	13.0		15.6	
7.	Decision making power	71.5	77.5**	2.87	74.0	1.11
		19.2	13.6		14.5	
8.	Life control	66.7	77.4***	4.23	75.5**	3.58
		21.3	14.8		16.3	
9.	Security	82.6	84.6	1.89	84.8	1.02
		17.7	10.9		13.7	
10.	Working ability	72.0	82.1***	5.74	85.9***	7.09
		18.3	17.0		12.0	
11.	Physical condition	63.7	67.8**	2.67	66.1	0.86
		20.7	18.1		19.3	
12.	Social relationships	75.7	81.1***	4.23	82.3*	2.64
		18.4	13.9		13.8	
13.	Activity and hobbies	52.5	60.2**	3.32	62.6*	2.68
		23.4	21.4		24.0	
Total SWB		59,8	73.6***	9.81	74.4***	7.68
		13.91	9.81		10.3	

---

* p<.01, ** p<.001, *** p<.0001; in scales 12 and 13 the number of comparison group members was 110. Total SWB includes scales 1-5, 8 and 10.

The level of the comparison group has been reached. The only statistically significant difference was in anxiety (t= 3.37, p<.001). Figure 7 shows the individual changes during the therapy on life satisfaction. Some of the changes were quite dramatic. When change is 50 points, life can look completely different. There were six clients whose life satisfaction dropped a little, but a large majority gained on the average 18 points. The figure 7 shows that the changes were highly variable. The change did not correlate with the length of therapy. Of course those who originally had low values gained most. Total SWB changes were also positive but similarly variable (figure 8). Changes were not as large as in individual scales, because there were some scales where the original level of SWB was already high.

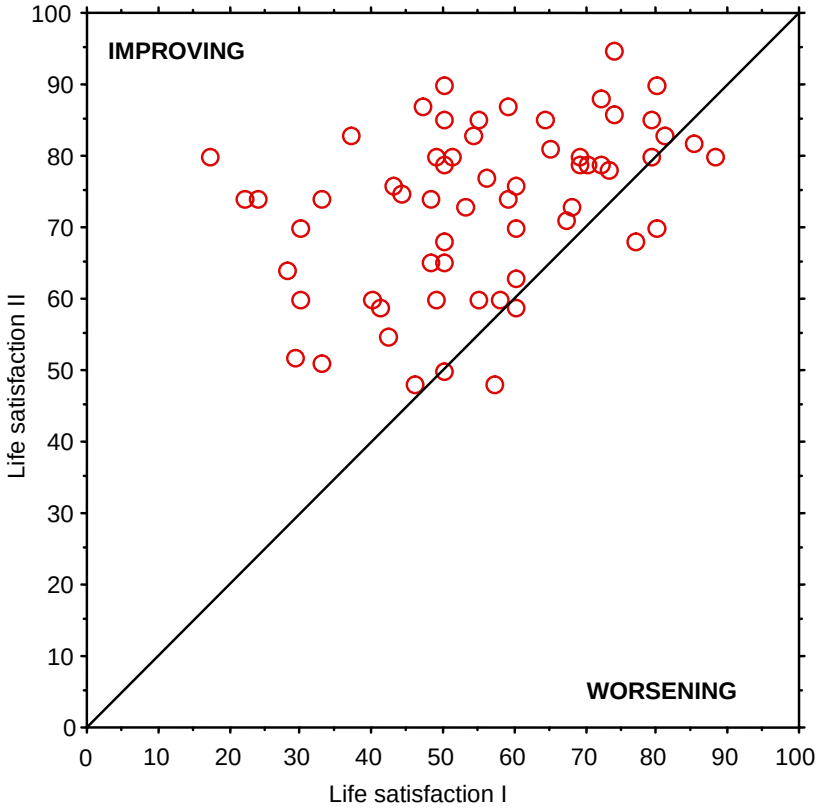


Figure 7. The scatter diagram of life satisfaction before and after therapy.

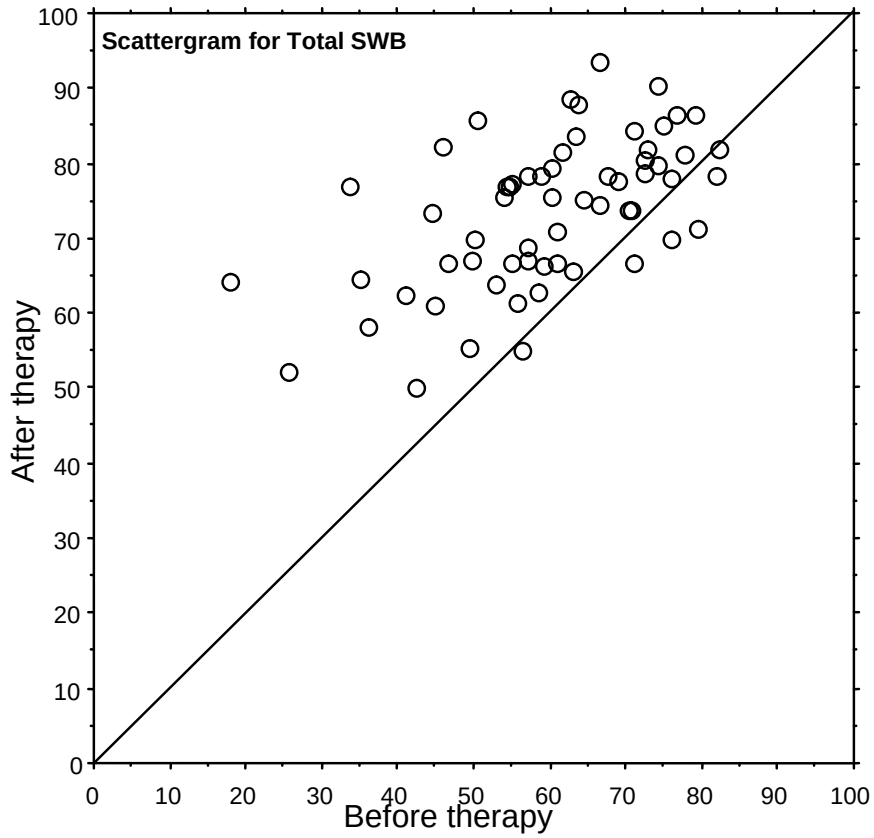


Figure 8. The scatter diagram of total Subjective Well-Being before and after therapy.

Problem experience. The changes in problem experience are shown in table 7. The changes were generally quite large. Before therapy problems were felt to be on the negative side of the scale, but after therapy all means were clearly positive. Figure 9 shows the large variance of changes. Only two changes were negative.

Table 7. The changes in problem experience during the therapy. The means, standard deviations and t-test differences (N=62)

	Before mean/sd	After mean/sd	t-test	p
1. Disturbing	1.69	6.95	15.56	.0001
	1.59	2.36		
2. Continuous	4.39	8.02	8.71	.0001
	3.05	2.40		
3. Guilt-inducing	4.50	7.28	6.94	.0001
	3.55	3.04		
4. Hinders work	3.94	7.03	6.87	.0001
	3.42	3.05		
5. Hinders social relations	3.74	7.05	7.05	.0001
	3.14	2.81		
6. Helplessness	3.42	6.92	7.68	.0001
	2.66	2.68		
7. Hopelessness	4.21	7.89	7.72	.0001
	3.04	2.48		
8. Thinking always	3.24	6.04	9.14	.0001
	2.13	2.23		
9. Worsening	5.18	8.37	10.29	.0001
	2.32	1.71		
Total problem exp.	4.39	7.32	14.22	.0001
	1.68	1.33		

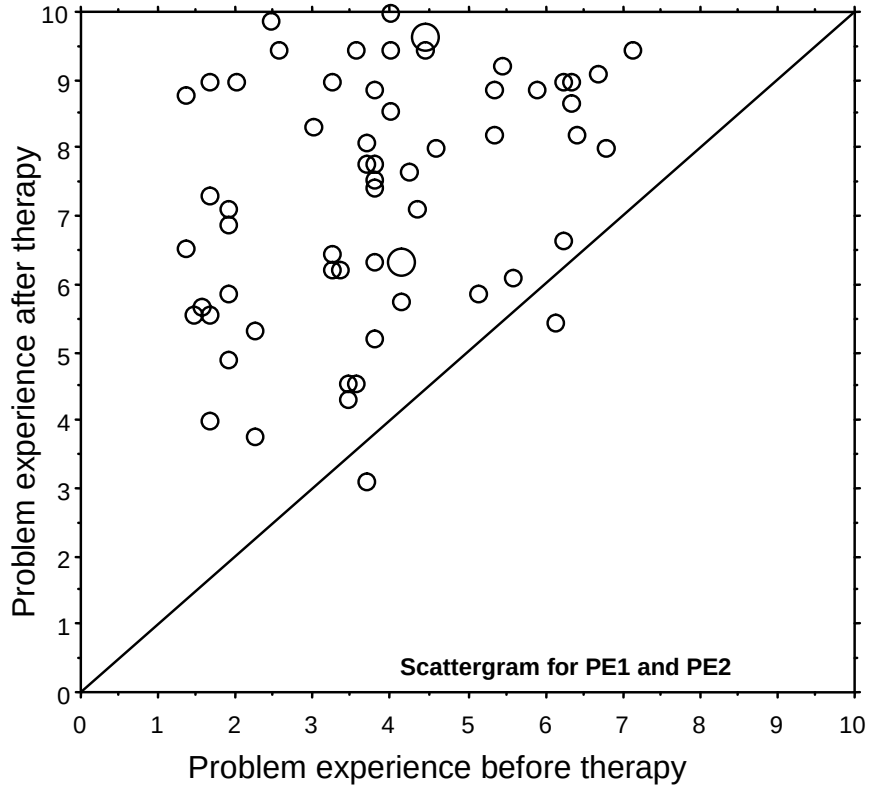


Figure 9. The scatter diagram of Total Problem Experience before and after therapy.

Table 8. The means, standard deviations and t-tests of therapist DVAS scales before and after therapy of the client group (60)

	Before mean/sd	After mean/sd	t-test	p
1. Anxiety	56.5	20.4	13.30	.0001
	20.8	15.6		
2. Anger	23.3	10.4	5.60	.0001
	19.3	13.1		
3. Mood	43.6	75.9	12.00	.0001
	19.6	13.5		
4. Vigour	55.8	78.0	9.24	.0001
	21.8	11.8		
5. Relaxation	41.6	73.5	13.90	.0001
	17.3	13.4		
6. Working ability	70.6	87.0	6.93	.0001
	21.8	15.2		
Total TWB	55.3	80.6	13.29	.0001
	16.5	11.4		

The therapists saw very large changes in their clients (table 8). This can also be seen in figure 10. There were a few increases in anxiety, but those clients did not have high levels originally. There were only three clients after therapy whose anxiety was rated to be above normal level (40).

However, there were 24 clients whose own ratings of anxiety were higher than 40 after therapy. The total TWB changes are shown in figure 11. Without a few exceptions these changes were positive.

The therapist ratings were originally somewhat more negative than those of the clients. The corresponding client ratings of anxiety, mood and working ability were 53.2, 56.0 and 72.0. The client and therapist ratings did not correlate highly. Correlations were .30 (anxiety), .44 (mood) and .51 (working ability).

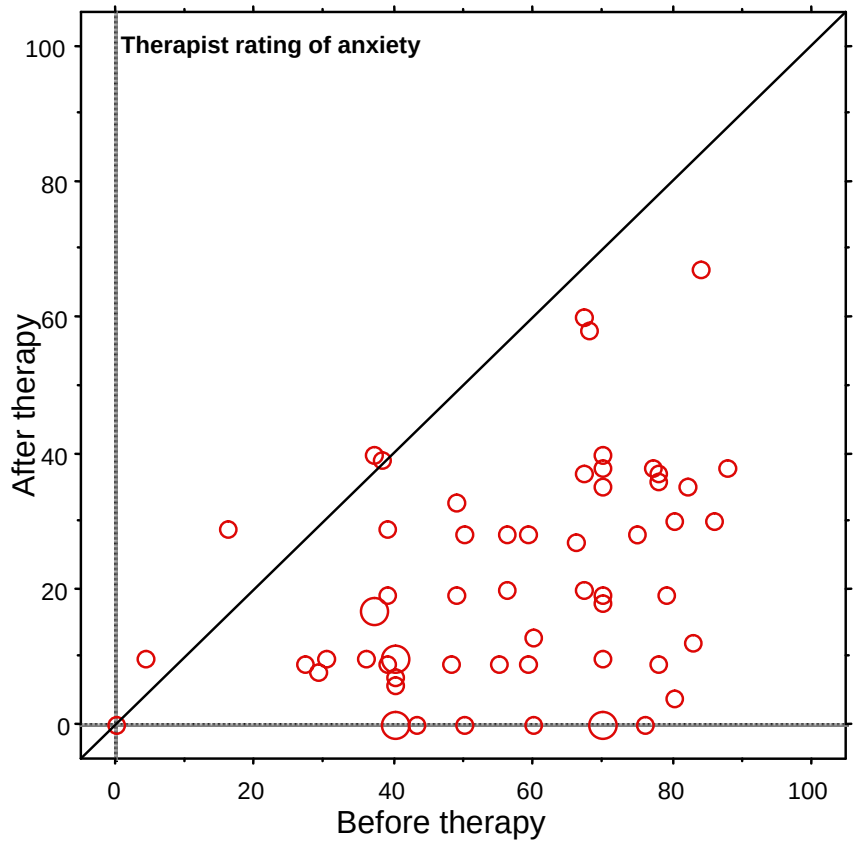


Figure 10. Scatter diagram of anxiety rated by therapists.

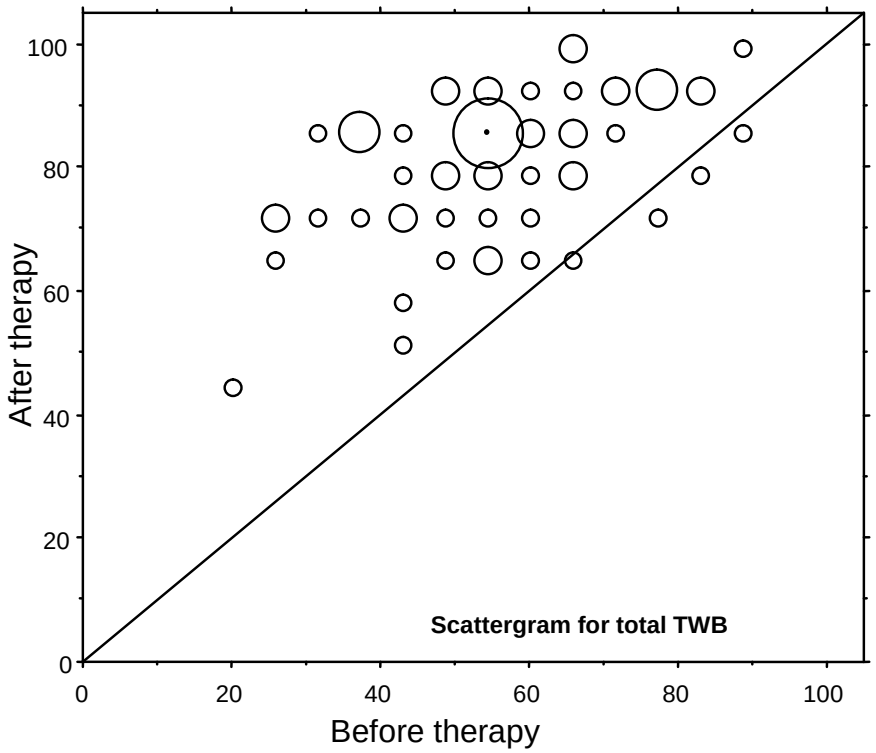


Figure 11. Scatter diagram of total Well-Being rated by therapists.

Experience of change. Both the clients and the therapists rated the changes during the therapy on 11 variables. The means and standard deviations are in table 9. Both the client and therapist ratings of change were highly positive. The differences between variables were quite small, though client rated problem change and personal growth highest. For therapists the highest changes were in problem and anxiety. Positive changes in adaptation to stress and in relationships were somewhat smaller. The client and therapist ratings corresponded fairly well. The average correlation of variables was .51 and total changes correlated .72. Two examples of client distributions are shown in figures 12 and 13.

Table 9. Ratings of change during the therapy. The means and standard deviations of therapists and clients.

	Clients		Therapists	
	Mean	sd	Mean	sd
1. Problem	8.39	1.31	8.18	1.24
2. Self-confidence	7.90	1.30	8.05	1.05
3. Functional ability	7.90	1.34	8.03	1.38
4. Relationships	7.48	1.61	6.98	1.53
5. Stress adaptation	7.45	1.48	7.52	1.22
6. Satisfaction	7.93	1.48	7.84	1.33
7. Growth	8.37	1.19	7.73	1.20
8. Mood	8.15	1.40	7.79	1.38
9. Anxiety	8.11	1.34	8.06	1.21
10. Life control	8.00	1.55	7.77	1.08
11. Total life situation	7.93	1.48	7.56	1.24
12. Total mean change	7.95	1.17	7.77	1.02

Scale: 0 = large negative change ... 5 = no change ... 10 = large positive change

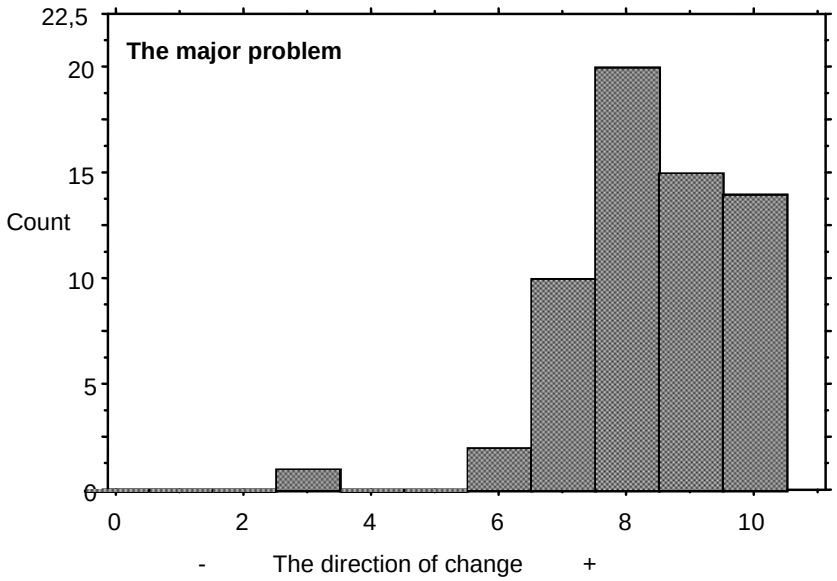


Figure 12. The retrospective change experienced by clients after therapy: The distribution of the major problem presented.

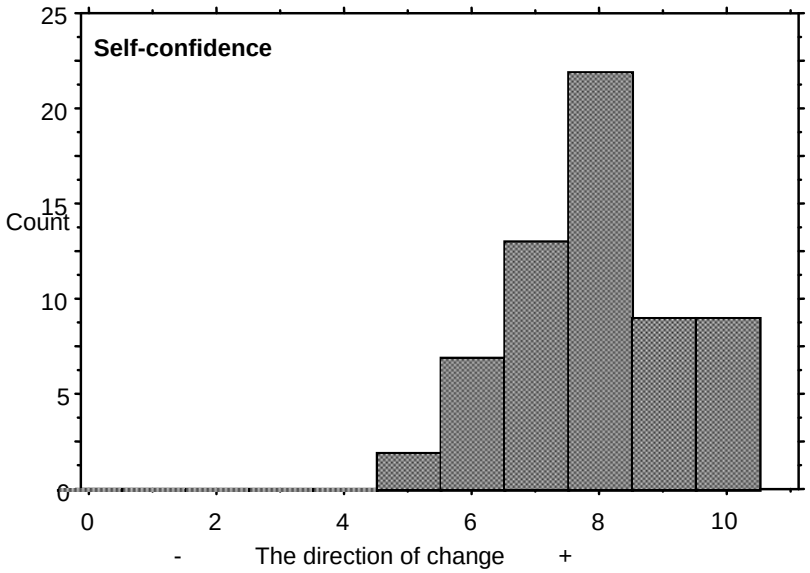


Figure 13. The retrospective change experienced by clients after therapy: The distribution of the self-confidence.

Correlations between the criteria. SWB and PE correlated .64, and TWB and PE .45 at pre-therapy evaluation. SWB and TWB correlate .33 which is not as high as could be expected. Thus therapists made their evaluations quite independently of client ratings. When PE was correlated with specific SWB scales, the highest correlations were with personal satisfaction (.64), mood (.51), well-being (.47) and anxiety (-.46). Subjective well-being evaluations are thus quite strongly related with problem experience.

Correlations of change scores. Five change scores could be extracted from the data. Three of these were real change scores (difference of before and after measures):
1. Total client SWB difference (7 scales)
2. Total problem experience difference (9 scales)
3. Total therapist TWB difference (5 scales)
In addition to these there were two subjective change scores:
4. Experienced change after therapy by clients (11 scales)
5. Change evaluation after therapy by therapists (11 scales)

The correlations of these five variables are shown in table 10.

Table 10. The correlation of real and subjective change scores after therapy

	1	2	3	4	5
1. Therapist TWB difference		.39	.27	.02	.17
2. Client SWB difference			.53	.32	.28
3. Problem experience difference				.44	.44
4. Experienced change (clients)					.71
5. Change evaluation (therapists)					

The correlations in table 10 show that different methods evaluating change did correlate, but the correlations were not high. An exception was that clients and therapists seem to agree after therapy about changes retrospectively.

It is also possible to compare the real changes and rated changes on similar variables (table 11). The first column shows that subjective experience of change did not correspond highly with pre-post -differences. On the other hand, the correlation of subjective change was quite high with SWB rating made at the same time. It can be concluded that ratings of change reflect more the present situation than the real change.

Table 11. The correlations of change scores and experienced changes after therapy.

	Change in SWB -scale vs. rating of change after therapy	SWB rating after therapy vs rating of change after therapy
1. Self-confidence	.18	.37
2. Relationships	.26	.37
3. Life satisfaction	.36	.29
4. Mood	.14	.40
5. Anxiety	.16	.59
6. Life control	.09	.26

Experienced change correlated quite highly with ratings of therapy (fulfillment of aims, alliance and satisfaction) and also with the after therapy measurements. The correlations with real change scores were lower.

Individual changes

The individual changes differed greatly (see figures below). The amount of change correlated strongly with before therapy measures. SWB1 correlated with SWBI-II -.72, PEI with PEI-II -.51 and TWBI with TWBI-II -.73. The worse well-being and the more serious problem, the larger changes were. Those whose general life situation was fairly good did not change much and similarly those who rated their problems not very difficult had not that much room for positive change.

Problem change. Clients made 9 ratings about their problems at each measuring point. The scale varied from 0-10, where 0-4 were negative values, 5 neutral and 6-10 were on positive values. 46 clients had at least 5 negative values from the 9 possible. Of these clients 10 were fully without negative values on the scales ("problem-free") after therapy and at the final follow-up available (6 months or 2 years or at both). In addition there was a client who was problem-free after therapy and at the final follow-up, but who had three low (0-4) values at 6 month follow-up. Three clients improved markedly during the therapy, but were problem-free only at follow-ups. Thus by quite strict criteria 10 (16%) were totally problem-free after therapy. The percentage is 23, if those four clients are included. If those clients who had less than 5 low scales before therapy (all had at least one) are followed, too, 9 clients were problem-free after therapy and in the final follow-up they participated in. Four had minor set-back at six month follow-up. This would raise the full recovery to 37%. These calculations show how difficult it is to draw lines on improvement. It is also instructive to show the mean number of low scales at each measuring points. These were 5.69, 1.58, 2.50 and 1.31. The curious increase of in the number of problem scales can be seen at six month follow-up. Five clients (12%) were highly problematic at the final two-month follow-up (5 or more scales). 24 clients (57%) had none. Of these definitely three are such that begin to improve only during the follow-ups.

Clinical changes. An alternative way of describing the results is the formula by Jacobson & Truax (1991; see also Kazdin, 1994, 55), which shows which individual changes were statistically significant. The formula is
t=mpre-mpost/sdiff
The re-test reliability of SWB is .90, but for PE and TWB it is not known. We estimate these at .80. Alphas at the first measurement were .66 and .87. PE was not a homogenous measure. Its re-test reliability can be expected to be higher.

At .05 level 32 (52%) of the clients achieve clinical change in SWB, 49 (79%) in PE and 42 (68%) in TWB.

By using a stricter criterion of change, in addition a change to a normal range was required. For SWB norm groups are available. Change was defined to reach normal range, if it was within one sd of norm mean (74.4 - 10.3). Adding this criteria to above there were 44% (27) clinical changes into normal range.

In PE it was required that none of the nine individual scales had negative values. Adding this criterion produced 42% (26) clinical changes into normal range.

For TWB norm groups or other meaningful limits were not available. We choose as normal a half standard deviation from the after therapy mean. This gave 74 as a boundary value. 55% (34) had significant clinical change and improved over this limit.

Table 12 shows the clinically significant changes and changes to normal level. This table can be read in many ways. Only 6 clients had no clinical changes at all. Why there were no changes? We can describe these clients as follows:
#7. Both SWB and TWB were originally at a normal range and problems were not very difficult. Presented problems: Low self-esteem, sexual problems, problems with father, fears, lack of assertiveness.
#9. TWB low and SWB almost reached the normal level. Problems were not very difficult. All measures changed positively, but only a little. Depression, tiredness, fears.
#16. SWB and TWB were at normal level and problems were not very serious. Public speaking anxiety.
#25 Both SWB and TWB were low, and problems were quite severe. Minor positive changes in each. Eating problems.
#45 SWB and TWB on normal range, problems not severe. No changes. Co-dependence, fear of failure, depression, anxiety. Many problems were mentioned in interview.
#46 SWB and TWB were somewhat below normal range, and problems severe. SWB and TWB changed but did not reach the criterion. Problems did not change at all. Too much energy, alcohol problems, anxiety, fears.

It is difficult to summarize these clients, but it seems that most of them gave quite high SWB ratings although their problems were quite severe if their verbal descriptions are observed.

Table 12. The clinically significant changes of clients and the changes which in addition attain normal level.

Clinically significant changes

		%
None	6	10
Only TWB	4	6
Only PE	7	11
Only SWB	2	3
SWB + PE	5	8
SWB + TWB	1	2
PE + TWB	13	21
SWB + PE + TWB	24	39
SWB	32	52
PE	49	79
TWB	42	68

Clinically significant changes + normal level

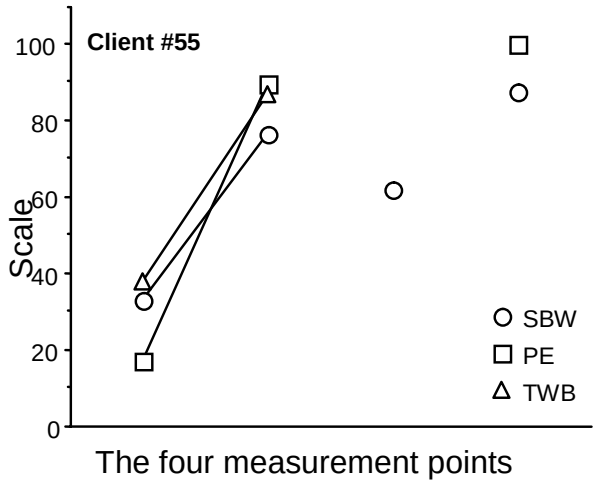
		%
None	13	21
Only TWB	10	16
Only PE	6	10
Only SWB	5	8
SWB + PE	4	6
SWB + TWB	8	13
PE + TWB	6	10
SWB + PE + TWB	10	16
SWB	27	44

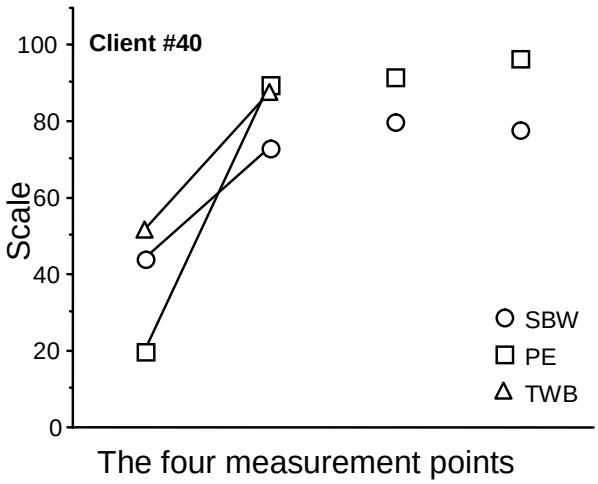
PE	26	42
TWB	34	55

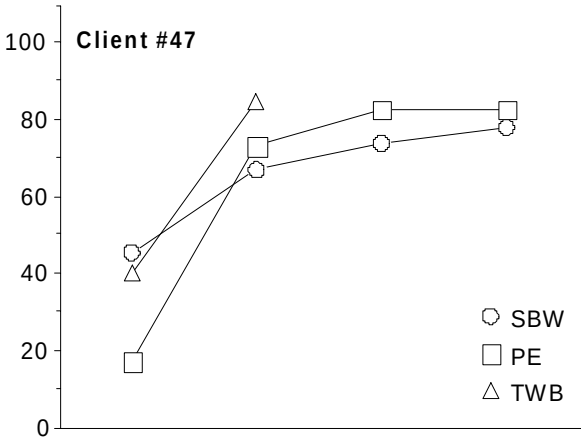
"Quantum change." Beside those who have not changed it is instructive to look for very large changes. We begin by looking those clients who did not have any problem experience scales below 5. There were 10 very large and stable changes (see figure 14). Of these 8 had marked and stable positive changes in SWB and/or TWB. Their self-described problems were as follows:

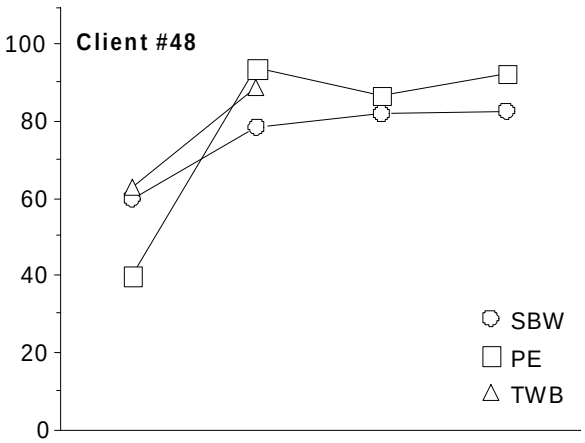
#2. Bad feeling, nothing interests me, alcohol and marital problems, violence.
#3. Mental health, marital problems, money problems
#10. Fear of dentist.
#40. Problems with work
#47. Study problems, problems in concentration and decision making
#48. Red cheeks and feelings of guilt
#51. Public speaking and social anxiety
#55. Tight nerves, physical symptoms, negative attitude to life

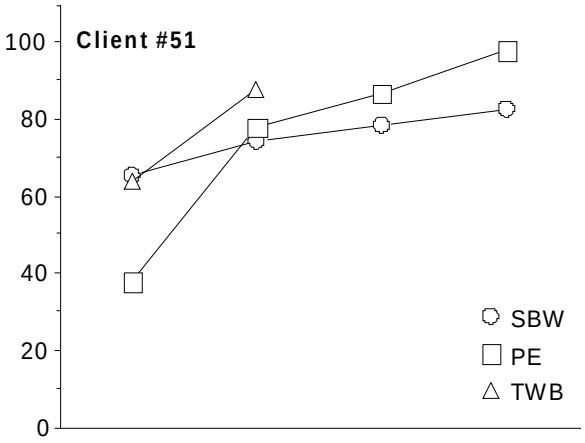
The greatest changes were in clients with code numbers 40, 47 and 55. These seem to be related to clients' attitudes or cognitions. It is interesting that fear of dentist can be rated as so severe. The profiles of these 8 clients are very different. They do not have much common with each others.

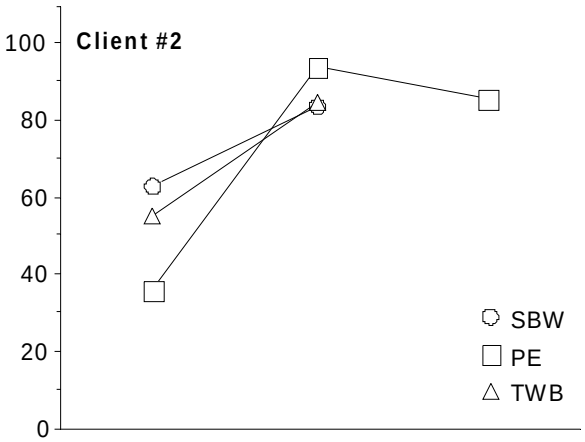


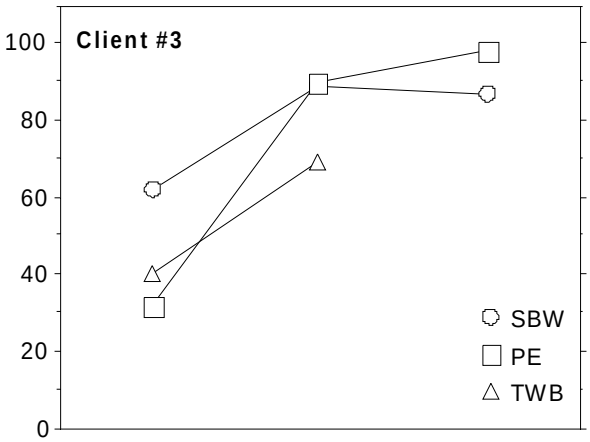

Client #40
100
80
60
40
20
0
Scale
SBW
PE
TWB
The four measurement points


Client #47
100
80
60
40
20
0
SBW
PE
TWB


Client #48
100
80
60
40
20
0
SBW
PE
TWB


Client #51
100
80
60
40
20
0
SBW
PE
TWB


Client #2
100
80
60
40
20
0
SBW
PE
TWB


Client #3
100
80
60
40
20
0
SBW
PE
TWB

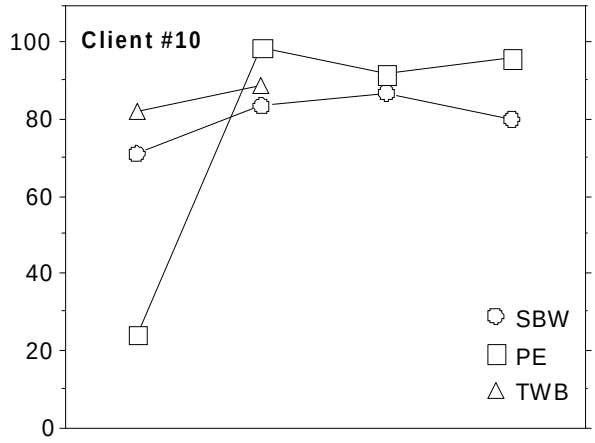


Figure 14. Clients having "Quantum change". Subjective Well-Being (SWB), Problem Experience (PE) and Therapist Well-Being (TWB) values at four measurement points. The scale of PE was multiplied by 10. In some cases the third or fourth measure may be lacking.

Expectations. Expectations of clients were measured by the difference of the present condition and expected change after therapy. This difference was on an average xx units on 0-100 -scale. Thus clients were quite optimistic about the changes during the therapy. This difference correlated from -.01 to .17 to client outcome differences (II-I) but somewhat higher with client true changes (.23) and rated changes (.26).

Follow-ups

The study included a six-month and two-year follow-ups. In the following the results are shown in graphs for all four measuring points. The graphs include those 42 persons about which we have follow-up data. When those 20 who did not stay with the study were compared with those 42 who did stay, there was a trend for those dropping out to be in worse condition. However, only the therapist TWB rating approached statistical significance (.054). For instance, pre-post change in SWB was very similar in both groups. There were no differences based on ratings of satisfaction toward therapy.

Figure 15 shows that the effects of therapy decreased at 6 months, but were back at 2 years. The reasons for this are not known. All three measures showed a similar trend. The decrease was statistically significant in all measures (pairwise t-test).

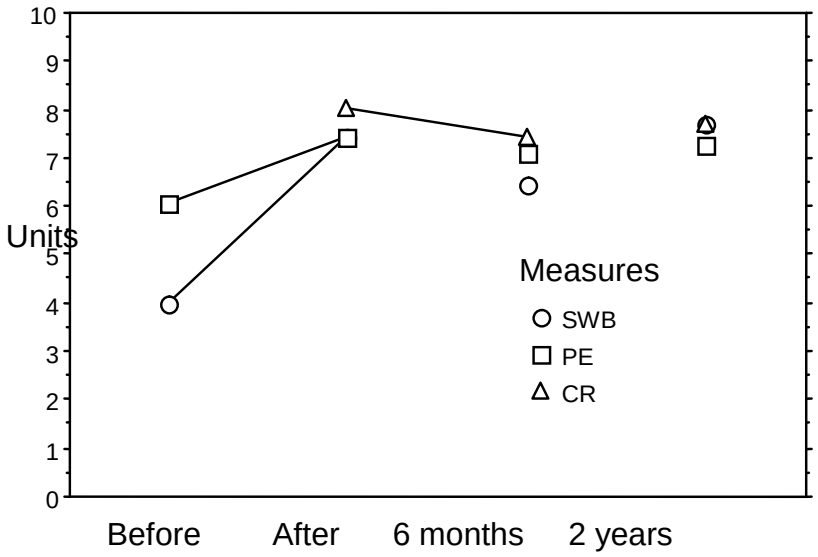


Table 13. Repeated measures analysis of variance over the four measuring points on Total Subjective Well-Being, Total Problem Experience, and Total Change Rating. Means, standard deviations and ANOVA tables.

Total Subjective Well-Being

Measures	N	Mean	Std. Dev.	Std. Error
SWB/I	41	60.95	13.70	2.14
SWB/II	41	74.13	9.61	1.50
SWB/III	41	71.01	12.56	1.96
SWB/IV	41	72.87	14.60	2.28

Source	df	S S	M S	F-test	P value
Between subjects	40	15172.12	379.30	3.05	.0001
Within subjects	123	15295.78	124.36		
treatments	3	4425.68	1475.23	16.29	.0001
residual	120	10870.10	90.58		
Total	163	30467.90			

Total Problem Experience

Measures	N	Mean	Std. Dev.	Std. Error
PEI	40	3.99	1.56	.25
PEII	40	7.44	1.61	.25
PEIII	40	6.44	2.13	.34
PEIV	40	7.71	2.11	.33

Source	df	S S	M S	F-test	P value
Between subjects	39	284.95	7.31	1.45	.0668
Within subjects	120	605.79	5.05		
treatments	3	344.89	114.96	51.55	.0001
residual	117	260.90	2.23		

Total	159	890.73			

Total Change Rating

Measures	N	Mean	Std. Dev.	Std. Error
CR/II	36	8.04	1.24	.21
CR/III	36	7.45	1.37	.23
CR/IV	36	7.68	1.65	.27

Source	df	S S	M S	F-test	P value
Between subjects	35	147.16	4.20	4.11	.0001
Within subjects	72	73.74	1.02		
treatments	2	6.45	3.22	3.35	.0406
residual	70	67.29	.96		
Total	107	220.9			

Other follow-up data

In table 14 responses about employment, help seeking and medication are described. The results were positive in the sense that only one person was without work at 2-year follow-up. 11 have sought help, but the clear majority had not. The question about medication must have been unclear at follow-up, because only 30 clients gave a response to it. The question concerned about maintenance medication. Perhaps some had short-term medication, but did not find place to mention it.

Table 14. Client unemployment, help seeking and medication before therapy and at two-year follow-up (N=42).

	Before therapy		2 year follow-up	
1. Unemployed	7	(16%)	1	(2%)
2. Sought help	21	(51%)	11	(27%)
3. Medication	10	(24%)	6	(14-20%)

Two-year follow-up included a question "How satisfied you are with your therapy after two years have passed?". The scale was 0=very dissatisfied ... 10=very satisfied. Mean was 8.54 and sd 1.79. In addition the clients was asked about their own input into therapy. Mean was 7.98 on a scale 0=not at all ... 10=very much.

Effect Strengths

The effect strength was calculated by dividing the mean pre-post difference by the before measure standard deviation. ES were as follows: SWB 0.99, therapist TWB 1.53 and PE 1.74. To control within-group design 62% can be taken out of the outcomes, though this correction may be too large. After this ES would be 0.61, 0.95 and 1.08. The ES of total SWB may be low compared to typical ES values, but this measure was very comprehensive and does not change as easily as specific measures. The SE of therapist TWB ratings may be suspect as a reactive measure, but SE of problem experience was a high one. If pooled sd's are used (see Cohen, 1988), uncorrected ESs were even higher: 1.12, 1.93 and 1.78.

But are these effects simply placebo effects or nonspecific effects similar to any eclectic therapy? The amount of placebo factors was estimated to be .40 in SE. The effect of NLP seems real, but could be placebo + common factors. On the other hand, on PE measure ES was considerably higher. This could be specific variance related to NLP or common factor variance can be observed best just on these kind of cognitive variables.

Comparison groups

Are the effects related to therapy? The major problem of the study is the lack of control group. Can the results be related to therapy or are they just "spontaneous" changes? Those who have before therapy sought help probably suffer quite stable problems. Their results were as good as those of first-timers on all measures. Similarly those on medication had as good an outcome as those not having medication.

On the other hand, regression toward mean could be an explanation. The amount of change was strongly related to the level of WB and problems. Those having low WB and difficult problems had generally large changes in the follow-ups.

What is the amount of change when a person with a similar SWB deficit is measured twice? There is data available which can help in answering this question. This data were described above in the subjects section. The level of 60 is selected as a cutting point in this comparison. Those having SWB less than 61 are considered having a SWB deficit.

Table 15. The SWB changes in six other groups having different follow-up times. All these groups including NLP group had an original level of SWB below 61.

Group	Number	Time	Mean I sd I	Change I-II
NLP-therapy	29	4 mo	48.1 10.4	19.7
Students D	12	1 mo	49.8 10.3	3.5
Students C	25	2 mo	52.9 6.4	4.4
Unemployed F	15	2.5 mo	49.5 10.8	0.7
Students B	25	4 mo	50.1 8.1	5.7
Company X	17	1,5 yrs	52.4 4.4	9.4
Students A	25	4 yrs	52.8 6.0	11.1

There are two possible mechanism that can raise the level of SWB in the groups presented in table 15: regression towards the mean and "spontaneous recovery" The four-week re-test correlation of five SWB variables (90 students) was .92 There probably is some "spontaneous" change, which increases during the length of the follow-up, but at four years it was only about half of therapeutic

effects. These comparisons support the therapeutic effects of NLP. Perhaps 25-30% of the total effect is due to spontaneous recovery. This corresponds to the estimations presented in the introduction.

Discussion

General

The conclusions and implications of Wampold's analyses (2001, 203-231) are plausible. There is no doubt that psychotherapy works. Dance therapy (Ritter & Low, 1996), physical exercise (Petruzzello, Landers, Hatfield, Kubitz, & Salazar, 1991; Craft, & Landers, 1998) and transcendental meditation seem be working according to similar criteria, too. New psychotherapy studies based on medical model are not very useful. The mass of studies point to a clear direction: Different kind of therapies seem to work. The variance in outcome seems to be mainly common factors and therapist effects. Placebo, working alliance, allegiance and competence are the main factors that explain outcome variance.

However, there is one reservation. There may be one combination of outcome studies where specific factors can be shown. When the problem is highly specific (spider phobia), the measure is specific (approaching and touching spiders) and the therapeutic method is also congruent with these being specific, specific factors probably show. Therapeutic specifity means that clients are coached to approach the spiders. This is true to some degree when specific cognitive ideas are measured with corresponding measures and the therapeutic training concentrates on these cognitions. Wampold (101-2) rightly criticizes BDI as sole criterion measure in depression outcomes. It does correlate highly with SWB measures, but is not a perfect measure of it.

There is a strong trend in psychotherapy research which favors medical model -based clinical trials. Huge efforts have not produced much beside the general result that psychotherapy and various other therapies are effective. Presently those therapies that have been studied a lot have the edge and have been listed as empirically supported therapy (EST), but more and more therapies will be added on those lists.

Could Neuro-Lingvistic Programming be one of those EST's? Probably, because it does have similar strengths than other therapies. Their proponents have developed ideas that could be called theory and they use a lot of specific techniques. Psychoanalysis, gestalt therapy or humanistic therapy are not on any firmer basis. Their basic theories are very controversial, too. Their sole asset is that they are older. Cognitive-behavioral therapies are on a firmer ground in the sense that there is a huge amount of both basic and applied research on behavior, learning and cognitions. However, when applied to psychotherapy the usefulness of this research has been questioned (Erwin, 1978).

The situation seems to be that the older therapies grudgingly accept or at least stand each others. Their proponents read only their own books and journals and do not think much about other schools. However, they unite to defend their standing when something new is coming. Something like NLP offers a good common target, especially when its proponents do promise too much and do present exceptional, abstract ideas about human mind and its working. Which is just the same what the gurus of old schools did.

A big problem is the concept of psychotherapy. On can ask what is common with the following methods used in various therapies (see Bergin & Lambert, 1994, 163-4):

- free association
- paradoxical intention
- teaching skills
- meditation
- massage
- primal scream
- correcting false beliefs
- punishment

The general principle seems to be: If you accept what I include, I accept what you include as psychotherapy. Treatment of soul can not include all those things. If you also look what are listed as problems (see e.g. Stevens, Hynan & Allen, 2002, 280-1) one wonders even more. How one can compare snake phobia, delinquency, smoking, nail-biting, jailed juveniles, marital discord, depression, headache, out-of-seat behavior (whatever it is) under the rubric psychotherapy studies? This is preposterous. In actuality, the more clinical trial research moves toward specifity, the less it concerns psychotherapy. We would reserve this title for complex verbal methods used to treat complex problems. It is just right word for eclectic psychotherapy. Other therapists should use specific names that tell what they are doing. A large part of therapeutic techniques are rather social methods and teaching methods. Most of cognitive-behavioral methods really concern teaching. That is their strength, but the results must not be compared with psychotherapy. In our opinion, much of psychotherapy should be replaced with clearly stated teaching strategies. Thus we need schools for living and not therapy based on medical model.

There is much in NLP which fits this idea, but they, too, want to sit on two chairs like everybody else. They want to train and educate the general public, but they also want to be psychotherapists. A very large part of the NLP courses has been offered for "bettering your life style and living skill". Most or the old schools are very restrictive in what they offer and sometimes even secretive, but NLP practitioners sell their product freely. At present the status of psychotherapy is so good that all kinds of therapies want to adopt that title.

Methodological problems

The medical model based clinical trial is an inviting one. It seems to be working well on agriculture, where the soil does not have expectations and where researcher's beliefs do not have effect on plants. In medicine problems are much greater. Medications can be studied, but only when ethically questionable placebo studies are used. In psychotherapy much of this breaks down. The idea of placebo can be discarded, because it cannot be meaningfully defined. Comparison of control and therapy groups is ethically questionable, because it is very well known that all kinds of therapies do work. Only if the problem is very specific and does not have a great effect on SWB, this kind of study is acceptable. Sometimes poor resources make the wait list controls possible. What is left is a comparison of two therapies which are probably useful. It is extremely difficult to design such kind of study. There will always be some kind of technical or methodological problems. There must be a very large number of clients, no researcher allegiance, similar clients and therapists in each group, etc. The problems of efficacy research are well summarized by Seligman (1995) and Wampold (2001).

An alternative could be that we collect a very large body of base rate data and use it as a basis for comparisons. This data body should be so large that all kinds of divisions and groupings could be formed on the basis of it. The groups studied in psychotherapy research are not necessarily a good sample of real clients. We should know what happens to clients having various problems in the

community. Any new method should produce better results than this. With this approach much more powerful within-group statistical analyses could be used. This kind of approach is not completely new. Sperry, Brill, Howard and Grissom (1996) describe large patient and community samples and compare outcomes of individual clients to this data. They also explicate a comprehensive evaluation package which includes SWB, self-rated symptoms, life functioning and mental health index. Curiously this is what the early researchers tried to do when they talked about spontaneous recovery (See Smith, Glass, & Miller, 1980, 10-18). Spontaneous recovery of neurotics was originally estimated to be very high, 30-70%.

Sperry, Brill, Howard and Grissom (1996) proposed a phase model of psychotherapy which consists of three phases or levels of psychotherapeutic efficacy. To these a fourth may be added, called "reorganization". Reorganization can also be called "quantum change" (Miller & C'de Baca, 2001). This kind of change concerns the whole personality, especially central life values. Changes are often lasting ones. Thus the four levels are:
1. Remoralization, emotional changes (hope)
2. Remediation, body-behavior changes (symptom relief)
3. Rehabilitation, skill-cognition changes (better skills or interpretations)
4. Reorganization, value-meaning changes (better life story or philosophy)

We modify these somewhat. The following model is based on problems. The targets of intervention can be:
1. Demoralization, dissatisfaction, low mood
2. Specific fears and negative habits
3. Multiple symptoms and problems
4. Negative attitudes and cognitions
5. Alienation, meaninglessness, unhappiness

The major therapeutic methods can be named as
1. Assurance, support
2. Relaxation and rest
3. Behavioral training/modification
4. Rhetoric, cognitive restructuring
5. Life skills teaching

Assurance is the "placebo effect", it is what any well-meaning person can do to help someone needing assistance. It helps especially in demoralization and is a component in any helping alliance. Relaxation aims to what the title says: to get a person to stop running, rest and relax. There are advanced methods that really put the body in peace. Behavioral training may be most useful in specific fears and negative habits, but it can be useful in teaching positive cognitions and reducing symptoms. Rhetoric aims to changing of opinions, attitudes and beliefs. Rhetorician shows that the present cognitions are not what they should be. Life skills teaching aims to showing ways what makes life happy and meaningful. The end point is a person whose identity is coherent, altruistic and recognizes one's self-worth.

NLP may be sitting on all five chairs. It tries include all therapeutic methods by using the 5 different levels of brain processing developed by Dilts (1990, 1). These levels are environment, behaviors, capabilities, beliefs and values, and identity. In NLP all these levels are important and although emphasis may be on a particular level, all must be taken care.

What then is psychotherapy? Perhaps the basic problem is that we are using this old word in order to cover anything that another person can do to another by verbal and even behavioral means.

Psychotherapeutic schools use too heterogenous methods that they could be compared with each others. Using manuals or comparing techniques is not a solution. The number of techniques is so large that comparisons became meaningless. The solution could be a new classification of helping methods along the lines presented above. There is not much evidence that any psychotherapy is more suitable to a particular group than an other one. An exception may be behavioral training program for specific fears and habits, but that is not really psychotherapy. This kind of teaching methods can be learned in a couple of hours and do not require long psychotherapy training. This is not to devalue this kind of training, because they are needed at least as much as psychotherapy.

There probably are client variables that are related to outcome (Garfield, 1994), but interactions of client and therapy factors are difficult to find (However, see as an example Barber and Monitz, 1996). Beutler (1979), Beutler et al.(1991) has developed hypotheses about symptoms to which particular therapies fit best. He thought that behavioral methods apply best to circumscribed, specific problems and also to clients whose defensive style is external and reactance low. We regard this as a plausible hypothesis.

Psychotherapy can thus mean practically anything. There are hundreds of talking and relating methods that go under this rubric. There is a rush to show that this particular method is also working, but what is not working? This is now much more interesting question. By looking SEs of individual comparisons on can find zero or near zero SEs, but the target groups are often special, like post-operative pain or no diagnosis groups in Stevens, Hynan and Allen (2000) study. Wampold (2001) argues that these zero results are completely random.

We do not believe that everything always works. Therapeutic alliance may not develop or expectations of clients may not met at all. In those conditions results may be meager. More interesting is the total incongruence of problems and methods. We predict that congruence is best at diagonal or near to it (figure 15). The reason why therapy results are generally good is that usually method and problem are congruent. With clients having spider phobia discussions of meaning are perhaps not helpful and not even started in such cases. Similarly going to very specific symptoms may not be the best way to start when a client feels hopeless and depressed.

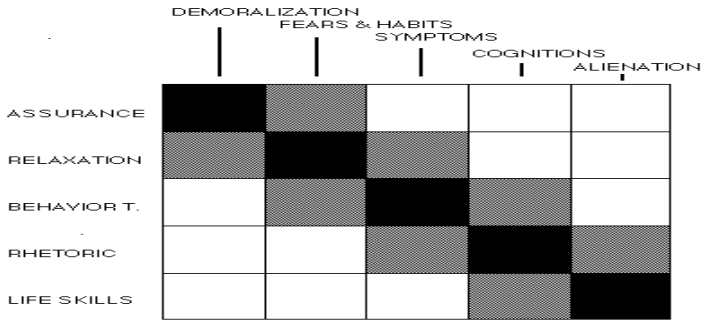


Figure 16. The congruence of problems and methods. Darkened areas are congruent, light ones are not congruent.

Secondly, problems of the clients are very often complex. The NLP idea about levels is probably a right one. However, complexity means that some problems and methods are bound to be congruent. Although the therapist always uses rhetoric, part of the problems can nearly always be described at that level. The more complex the problems are, the more question is about meaning and purpose in life. Thus those therapies that emphasize this level may be just right fore very many clients.

Thirdly, the differences between different approaches are not that great. All therapist use assurance, perhaps relaxation and rhetoric, too. For this reason it is understandable that placebos work as well as they do. Similarly therapy given by motivated layman works very well.

Those who argue for some kind of new therapy will practically always succeed in showing that their therapy is effective. Their evidence can be denied because of problems based on clinical trial model, but if they persist, the results will be positive. To show that it works better than other therapies is quite hopeless task. If the problem is very specific, psychotherapy is not needed, but behavioral training, and if it is not specific, anything helps in it. In specific problems the question is about learning and not so much therapy. It strains our pride to admit that even the most esoteric therapies may work about as well as our "scientific" therapies.

Validity of the measures

Though the validity of self-report measures are often criticized (Paul, 1986, 44), they are commonly used in outcome studies. Of course the more severe the problems are, the more difficult evaluations of own behaviors and traits are. Other alternatives are therapist ratings and use of assessors. Both the clients and therapists have much on the stake and thus there may be a positivity bias in the results. This is probably even more true of therapists than clients. Assessors may be an alternative, but often it is not a practical way. They can observe the clients only on a certain occasions which may not be typical of client behavior. Often they know which group they study or can deduce the membership of a certain person.

Particularly those persons having depression or anxiety problems are the ones who really know their situation. This means that there is no alternative to self-report data. Luckily various measures do give similar results, though there are some minor differences. Miller and Berman (1983) found mean SE of self-report marginally higher than mean observer SE. Matt and Navarro (1997) state that "Highly reactive outcome measures - as defined by Smith et al. (1980) - have been associated with higher effect sizes in 6 of the 9 meta-analyses reporting findings on these relationships" (21).

The particular selection of methods needs comments. In our opinion good global measures of SWB do not exist. The nearest what we aim are simple scales of happiness and personal satisfaction (Sandvik, Diener, & Seidlitz, 1993). These scales should reflect widely positive aspects of life, like meaning, security, optimism, life control, etc. The present scales do correlate highly with older scales of happiness and satisfaction and they differentiate groups having various illnesses and adversities (Ojanen, 2000; 2002). The internal consistencies and retest correlations have also been high.

The problem experience method was specially tailored to the present study. It had an adequate internal consistency and it correlated with other criterion variables.

Effectiveness

The ratings of the clients about the therapy were very positive. The therapeutic relationships were highly positive, too. These kind of evaluations are not always thought to be important, but positive reaction is a necessary precondition to a successful therapy. Without it results are not possible.

The clients felt they changed into better. This kind of subjective result is very important, because it gives courage and hope. It would be interesting to see results where the clients changed better according to repeated measurements, but their experience was negative. Also, it would be interesting to look results where the real changes are nonexistent or even negative, but the experience of clients was highly positive. Actually, these kind of results are not very probable.

In practice both kinds of information lead into the same conclusion. Consumer reports on effectiveness of psychotherapy are very positive (Seligman, 1995) and all meta-analyses show psychotherapy to be efficient (Wampold, 2001). Of course, there must be something really wrong with therapy, if is not effective or efficient or both. Perhaps some kind of forced therapy might work, but would be unethical to offer. It may be possible to create a very good atmosphere without having any real results. We do not believe this is common.

All measured changes during NLP-based therapy were positive. Though it is possible that spontaneous recovery explains part of the results, this is not plausible for the following reasons:
1. The normal level of SWB was practically reached and the results lasted quite well.
2. The spontaneous recovery was much less on other groups having originally a low level of SWB.
3. The ES values were quite similar to other psychotherapy studies.

The results do not provide any light for the question of common vs. specific factors. In our opinion specific factors are possible in only special situations which was not true of this research. The problems of the clients were not specific but generally consisted of many symptoms and problems.

The results do not tell about the suitability of NLP to specific problem groups. It was observed that outcome correlated strongly with well-being deficit and problem severity. This is not a typical result (Garfield, 1994, 209-210). The reason may be that there simply is not that much room for change in SWB measures made according to directions of Cummins (1996). Only those who have very severe SWB deficit can change much. The situation is different with PE scales which are alike traditional measures. Pretest-change correlation was not as high in this measure and was only -.30 when those having severe problems were taken apart (N=47, mean PE<5).

The present study seems to be a first attempt to evaluate NLP. From the point of clinical trial it has many weaknesses, but its strength was that it was done in real-life conditions. We conclude that in this particular example NLP-based psychotherapy was effective and there are good reasons to believe that its was efficient, too.

We feel that there is no point in manualized rigid therapies, but instead those means are profitable that strengthen the alliance and reinforce social psychological mechanisms for the therapeutic relationship. This does not mean that therapists should be eclectic but they can use a wide collection of methods based on sufficiently coherent principles like cognitive-behavioral and psychodynamic systems. It seems quite plausible that Neuro-Lingvistic Programming can offer this kind of secure basis for therapeutic relationship.

References

Barber, J.P., & Muenz, L.R. (1996). The role of avoidance and obsessiveness in matching patients to cognitive and interpersonal psychotherapy: Empirical findings from the treatment for depression Collaborative Research Program. Journal of Consulting and Clinical Psychology, 64, 951-958.

Beutler, L. (1979). Toward specific psychological therapies for specific conditions. Journal of Consulting and Clinical Psychology, 47, 882-897.

Beutler, L.E., Mohr, D.C., Grawe, K., Engle, D., & MacDonald, R. (1991). Looking for differential treatment effects: Cross-cultural predictors of differential psychotherapy efficacy. Journal of Psychotherapy Integration, 1, 121-141.

Borcovec, T.D., & Ruscio, A.M. (2001). Psychotherapy for generalized anxiety disorder. Journal of Clinical Psychiatry, 2001, 62 [suppl 11], 37-42.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Craft, L.L., & Landers, D.M. (1998). The effect of exercise on clinical depression and depression resulting from mental illness. Journal of Sport & Exercise Psychology, 20, 339-357.

Dilts, R. (1990). Changing belief systems with NLP. Capitola, Cal: Meta Publications.

Cummins, R.A. (1996). The domains of life satisfaction: An attempt to order chaos. Social Indicators Research, 38, 303-328.

Erwin, E. (1978). Behavior therapy. Scientific, philosophical, & moral foundations. London: Cambridge University Press.

Eysenck, H. (1952).

Garfield, S.L. (1994). Research on client variables in psychotherapy. In A. E. Bergin & S.L. Garfield (Eds.), Handbook of psychotherapy and behavior change (4th ed.). New York: Wiley, pp. 190-228.

Greenberg, L., Elliot. R., & and Lietaer, G. (1994). Research on experiential psychotherapies. In A. E. Bergin & S.L. Garfield (Eds.), Handbook of psychotherapy and behavior change (4th ed.). New York: Wiley, pp. 509-539.

Hunter, J., & Schmidt, F. (1990). Methods of meta-analysis: Correcting error and bias in research findings. Newbury Park, CA: Sage.

Jacobson, N.S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology, 59, 12-19.

Kazdin, A.E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S.L. Garfield (Eds.), Handbook of psychotherapy and behavior change (4th ed.). New York: Wiley, pp. 19-71.

Kazdin, A.E. (2001). Almost clinically significant (p<.10): Current measures may only approach clinical significance. Clinical Psychology: Science and Practice, 8, 455-462.

Lambert, M.J., & Bergin, A.E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S.L. Garfield (Eds.), Handbook of psychotherapy and behavior change (4th ed.). New York: Wiley, pp.143-189..

Lipsey, M.W., & Wilson, D.B. (1993). The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. American Psychologist, 12, 1181-1209.

Matt , G.E., & Navarro, A.M. (1997). What meta-analyses have and have not taught us about psychotherapy effects: A review and future directions. Clinical Psychology Review, 17, 1-32.

Muurinaho, J. (2000). Mielenterveyskuntoutujien oma-apuryhmät. [Self-help groups of mental health rehabilitees.] Masters' thesis, department of psychology, University of Tampere.

Nietzel, M.T., Russell, R.L., Hemmings, K.A., & Gretter, M.L. (1987). Clinical significance of psychotherapy for unipolar depression: A meta-analytic approach to social comparison. Journal of Consulting and Clinical Psychology, 55, 156-161.

McDermut, W., Miller, I.W., & and Brown, R.A. (2001). The efficacy of group psychotherapy for depression: A meta-analysis and review of the empirical research.

Miller, R.C., & Berman, J.S. (1983). The efficacy of cognitive behavior therapies: A quantitative review of the research evidence. Psychological Bulletin, 94, 39-53. Clinical Psychology: Science and Practice, 8, 98-116.

O'Connor, J., & Seymour, J. (1993). Introducing NLP. Neuro-linguistic Programming. Rev. ed. London: Aquarian.

Ojanen, M. (2000). Effects of illness and adversity on quality of life. Teoksessa J.Jobin (ed.) Advances in cardiopulmonary rehabilitation, 198-210. Champaingn, Ill.: Human Kinetics Press.

Ojanen, M. (2001). Graafiset analogia-asteikot elämänlaadun ja hyvinvoinnin mittauksessa. Teoksessa S. Talo (toim.). Toimintakyky – viitekehyksestä arviointiin ja mittaamiseen. Turku: KELAn Sosiaali- ja terveysturvan katsauksia, 49, 207-225.

Ojanen, M., Tuori, T, & Lauren, H. (1997). The Effects of the Finnish National Schizophrenia Program (1981-87): Treatment and rehabilitation of long-stay Mental Patients, Evaluation, 3, 133-154.

Paul, G.L. (1986). The nature of DOC and QICS encoding devices. In G.L. Paul (ed.) Assessment in residential treatment settings. Principles and methods to support cost-effective quality operations. Champaign, Ill: Research Press, 63-122.

Petruzzello, S.J., Landers, D.M., Hatfield, B.D., Kubitz, K.A. & Salazar, W. (1991). A meta-analysis on the anxiety reducing effects of acute and chronic exercise: Outcomes and mechanisms. Sport Medicine, 11, 143-182.

Ritter, M.A., & Low, K.G. (1996). Effects of dance therapy: A meta-analysis. The Arts in Psychotherapy, 23, 249-260.

Sandvik, E., Diener, E. & Seidlitz, L. (1993). Subjective well-being: The convergence and stability of self-report and non-self-report measures. Journal of Personality, 61, 317-342.

Seligman, M.E.P. (1995). The effectiveness of psychotherapy. The consumer Reports Study. American Psychologist, 50, 965-974.

Smith, M.L., Glass, G.V., & Miller, T.I. (1980). The benefits of psychotherapy. Baltimore: The Johns Hopkins University Press.

Stevens, S.E., Hynan, M.T., & Allen, M. (2000). A meta-analysis of common factor and specific treatment effects across the outcome domains of the phase model of psychotherapy. Clinical Psychology: Science and Practice, 7, 273-290.

Taylor, S. (1996). Meta-analysis of cognitive-behavioral treatments for social phobia. Journal of Behavior Therapy & Experimental Psychiatry, 27, 1-9.

Wampold, B.E. (2001). The great psychotherapy debate. Models, Methods, and Findings. Mahwah, N.J.: Lawrence Erlbaum.

Öst, L. G. (1989). One-session treatment for specific phobias. Behavior Research and Therapy, 27, 1-7.