

Towards Alternative Criteria for the Validation of Psychological Treatments

Francisco Javier Carrascoso López^{*1} y Sonsoles Valdivia²

¹*Servicio Andaluz de Salud, Sevilla, España* ²*Universidad de Almería, España*

ABSTRACT

The Chambless Task Force criteria have become the standard for the research on outcomes of psychological treatments. However they present conceptual weaknesses and do not contemplate the vast complexity of the research and development of new psychological treatments. The current paper presents a critical review of such weaknesses to propose six sets of alternative criteria for the validation of psychological treatments, which are multidimensional and horizontal: Conceptual criteria, connections with basic research, research on processes of change, research on efficacy, research on effectiveness; and research on efficiency. The six sets of criteria emphasize the relation between the development of psychological treatments and the basic research on psychological processes, the research on the processes of change underlying a treatment, and the research on treatment efficacy, effectiveness and efficiency. Finally, future perspectives within our proposal are presented. *Keywords:* Change processes, Efficacy, Effectiveness, Efficiency, Treatment validation, Empirically validated treatments.

RESUMEN

Los criterios de la Comisión Chambless se han transformado en el estándar en la investigación sobre los resultados de los tratamientos psicológicos. Sin embargo resultan débiles conceptualmente y pueden hacer poca justicia a la enorme complejidad de la investigación y desarrollo de nuevos tratamientos psicológicos. El presente trabajo parte de una revisión crítica de las debilidades de los criterios de la Comisión Chambless para proponer seis conjuntos alternativos de criterios de validación de los tratamientos psicológicos, de naturaleza horizontal y multidimensional: conceptuales, conexiones del tratamiento con la investigación básica, investigación sobre procesos de cambio, investigación sobre la eficacia del tratamiento, investigación sobre la efectividad del tratamiento, e investigación sobre la eficiencia del tratamiento. Los seis conjuntos de criterios enfatizan la relación entre el desarrollo de los tratamientos psicológicos y la investigación básica sobre los procesos psicológicos, la investigación de los procesos de cambio presupuestos por el tratamiento, y la investigación sobre la eficacia, efectividad y eficiencia del tratamiento. Finalmente, se presentan líneas de discusión futuras.

Palabras clave: procesos de cambio, eficacia, efectividad, eficiencia, validación de tratamientos psicológicos, tratamientos empíricamente validados

*Correspondence concerning this article should be addressed to the first author: c/ Santa María de Ordás, 7, Bajo C, 41008 Sevilla, España. E-mail: carrascoso@andaluciajunta.es

The Chambless Task Force criteria (Chambless, Baker, Baucon, *et al.*, 1998; Chambless, Sanderson, Shoham, *et al.*, 1996; Task Force on Promotion and Dissemination of Psychological Procedures, 1995) have guided recent research on treatment outcomes, and have promoted several reviews of the available literature with very consistent results. Thus, they have proven remarkably successful as the standard to conclude whether a psychological treatment can be validated or not. Basically, the Chambless Task Force criteria specify the minimum empirical evidence required as well as the conditions in which such evidence must have been obtained, to consider a psychological treatment either validated or in the way towards its validation. These are, then, terminal criteria, since they are focused in the evidence already obtained, not in the *process* by which it is obtained.

The Chambless Task Force criteria have a clear operational and empirical character, and are restricted to the research on the psychological treatments efficacy. Thus, neither the treatment effectiveness or efficiency (despite the efforts by Chambless & Hollon, 1998; Chambless & Ollendick, 2001), nor the basic research on the change processes underlying the effects of a treatment, nor the basic research on the psychological processes involved, are contemplated. Consequently, Chambless Task Force criteria barely inform about the philosophical foundations and the model of psychopathology underlying a particular treatment, which makes difficult to conclude about the validity of such treatment. Instead, they only contemplate the minimal conditions that warrant the reliability of the efficacy data.

Despite the pragmatic purposes of these criteria, their emphasis in reliability and efficacy is not conceptually neutral and it is based on two aspects. First, a certain treatment is specific for a particular problem defined by its topography. Second, the technological aspect is key for the development of procedures for the behavioural change. Regarding the first aspect, the Chambless Task Force criteria adopt the basic rationale of the DSM when assuming that it is possible to identify efficacious treatments for particular problems. However, accepting this is dangerous from an historical, conceptual and empirical standpoint, as long as it implies that (Berrios, 1984): (a) we have already defined all possible symptoms with clinical and research purposes; (b) such definition is complete and comprises all necessary information; and (c) consequently, psychopathology can be identified by any observer, which means that the context in which such psychopathology is identified is just "noise" producing quantitative variability treatable as artefact with statistics. In other words, psychopathology exists independently from the observer, who has just to pay good attention in order to get a *reliable* definition. Contrarily, historical and empirical research has shown that the conceptual categories in psychopathology are constantly changing due to social, economical and professional variables, with important practical consequences (for instance, Berrios & Porter, 1995; Bourgeois, 1994; Callahan & Berrios, 2005; Goldstein, 2001; Pérez Álvarez, 2003).

Regarding the emphasis in the technical aspect of the development of procedures for the behavioural change, and given the focus in the reliability of the research on efficacy, the Chambless Task Force criteria disregards the fact that the language utilized in the validation process is medicine-imported, as Goldfried and Wolfe (1998) noted. This is essentially observed in the dependent variables employed (usually the reduction

in duration, frequency, intensity, etc. of a number of symptoms), and in the concept of “treatment” that is implicit.

The operational definition of the dependent variables employed is restricted to the description of discrete response topographies, and does not include the functional relations among variables. However, if behaviour is to be defined in functional terms, response topographies can never stand for psychological processes. In fact, one of the Chambless Task Force criteria paradoxes is that they cannot be applied, for instance, to the token economy because it is utilized for several behaviours topographically different to each other (Chambless *et al.*, 1998).

The emphasis in symptoms reduction, within such model of therapeutic change, as the key to validate a psychological treatment has several implications. First, variables like client’s quality of life or wellbeing, are not so much considered, hiding the fact that in the clinical practice, very different change patterns can be identified across clients, but all sharing the same function, this is, acting in the direction of a life committed to the personal values (Follette, Bach & Follete, 1993). In line with this, within the third wave of behavioural therapies (Hayes, 2004), it is stated that behaving in accordance to the personal values is a more sensitive measure of psychological wellbeing than any other (Hayes, Strosahl, Bunting, Twohig, & Wilson, 2004; Luciano, 2001; Páez-Blarrina, Gutiérrez-Martínez, Valdivia-Salas, & Luciano-Soriano, 2006). The exploration and systematic work on the clients’ values allows connecting the therapy with the change processes and better utilizing the available knowledge about psychological processes in the clinical context. Second, the emphasis in symptoms reduction prevents clinicians and researchers from connecting the basic research on psychological processes to the development of psychological treatments. Third, the Chambless Task Force criteria only require a number of well-controlled studies, carried out by independent investigators, with favorable results in terms of symptoms reduction (which can hardly be accepted as warranty of validity and conceptual relevance), to consider a given treatment as empirically supported. Thus, any non-medical intervention meeting such requirements can be validated as a treatment with empirical support, regardless of its vague conceptual foundations and relevance (Rosen & Davison, 2003). Fourth, the definition of “treatment” implicit in the research on treatment outcomes has some difficulties. Usually a psychological treatment is understood as a set of therapeutic techniques, without making explicit the change processes (for instance, acceptance vs. control) responsible for a particular improvement in the life of the individuals.

Actually, the manuals only describe the implementation of the techniques on the side of the therapist and the expected improvement on the side of the clients, without making explicit the change processes responsible for such an effect. This way, we could observe, for instance, an improvement in social skills that were never explicitly trained, without being able to explain where such improvement comes from (Wilson & Luciano, 2002). In this direction, a movement is being started from *empirically validated treatments to change principles with empirical support* (Rosen & Davison, 2003), or to *change strategies and theories* (Westen, Novotny, & Thompson-Brenner, 2004). We assert that an alternative definition of treatment which specified the relevant clinical behaviours of both the therapist and the client, and the connections between the development of

a psychological treatment and the basic research on psychological processes, is needed. Such definition would allow better testing the change processes and theories previously defined in operational terms.

The considerations just made might be interpreted as weaknesses that limit the heuristic potential of the Chambless Task Force criteria, their capacity to capture all the aspects of the development of psychological treatments, and their clinical and marketing usefulness. In an attempt to improve such weaknesses and to contemplate the complexity of the process for the development and implementation of the validation criteria, we present a series of alternative criteria which comprises:

1. The treatment's potential for innovation, and/or its conceptual relevance and connection with the available research about psychological processes.
2. The diverse aspects of the development of new change technologies, at both basic and outcomes research levels.
3. The conditions that facilitate the applicability and dissemination of a treatment.

CHARACTERISTICS AND USAGE OF THE ALTERNATIVE CRITERIA FOR THE VALIDATION OF PSYCHOLOGICAL TREATMENTS

We are proposing six alternative sets of criteria, which are: 1. Conceptual; 2. Connections with the basic research; 3. Research on processes of change; 4. Research on treatment efficacy; 5. Research on treatment effectiveness; and 6. Research on treatment efficiency. All sets of criteria are independent among each other, and address different, and interrelated, aspects of the literature on psychological treatments. For each set of criteria, we will distinguish between the well-established treatments, and the treatments in the way towards their validation. The operational definition of each set of criteria establishes the conditions that a treatment should meet in order to be considered as well established or in the way towards its validation in any of the areas under study.

Functional characteristics of the alternative criteria

The criteria we are proposing are horizontal, that is, there is neither a criterion more important than the others, nor a hierarchy of them. We consider that the development of psychological treatments is a very complex and continuous conceptual and empirical enterprise that can grow simultaneously in different directions interacting to each other and in different speeds. With these criteria we try to capture the continuous and bidirectional nature of the treatment validation process, along with all the aspects of its development and praxis, avoiding the risks derived from committing *exclusively* to outcomes research, or the conceptual coherence and clarity.

If a treatment does not meet all the criteria it is not automatically dismissed, but that information is used to detect the flaws of such treatment, and thus, the areas in need of further exploration at empirical and/or conceptual levels. The more criteria are met, the more established is a treatment and the more preferred as a choice option.

Usage of the alternative criteria

The application of the alternative criteria involves two steps. During the first step, that we will call *analytic*, each set of criteria should be applied individually to the literature on a given psychological treatment. As a result, we will obtain information about the degree in which each of the areas of a treatment is well established, in other words, this analytic step yields quantitative information about the state of the literature on a certain psychological treatment. This information is interesting and useful to researchers, to the extent that it may guide the establishment of a research plan on that treatment, but not necessarily to clinicians, politicians, and the general public. Again, the application of each of the six sets of criteria to the literature on a given psychological treatment will produce as a result whether that treatment is well established or in the way towards its validation, in each of the six sets of criteria individually considered.

During the second step, that we will call *synthetic*, the information obtained in the previous step will be properly elaborated so as to answer three questions about a given psychological treatment:

1. Which is the treatment's potential for conceptual, empirical and technological innovation? Is its potential for innovation larger, same, or smaller than other treatments' potential? The previous application regarding the conceptual, connections with basic research, and research on processes of change criteria, will allow answering those questions. The potential for innovation will be directly proportional to the number of relevant criteria for which the treatment is well established. If the treatment is well established in the three sets of criteria that define its potential for innovation, this will be larger than the potential for innovation of other treatments, which will be considered in the way towards their validation in one or more of those criteria.
2. Which is the treatment's clinical utility? Is its clinical utility larger, same or smaller than other treatments' utility? The previous application of the criteria regarding research on processes of change, and research on treatment efficacy and effectiveness will yield the answers. Again, the clinical utility will be directly proportional to the number of criteria for which a given treatment is well established. If a given treatment is well established in the three set of criteria mentioned before, its clinical utility will be larger than that of other treatments that do not meet the three sets of criteria.
3. Which is the treatment's potential for dissemination? Is its potential for dissemination larger, same or smaller than that of others? The previous application of the criteria regarding connections with basic research, research on processes of change, and research on treatment efficiency will answer such questions. Again, a directly proportional relation will be established between the number of criteria for which the treatment is well established and its potential for dissemination. Thus, if a treatment is well established in the three sets of criteria mentioned before, its potential for dissemination will be larger than that of other treatments that do not meet the three sets of criteria.

The application of this synthetic step yields qualitative information about the state of the literature on a given psychological treatment, which is important for researchers, and also for clinicians, politicians, and the general public. In other words, such information should influence the decision-making process that those professional undergone in their

daily work by placing certain treatments in a better position than others, and also, by informing about the area that research efforts should be allocated for the sake of the good progress in clinical psychology.

Examples of possible results after the application of the alternative criteria

Say a treatment is well established in relation to the conceptual criteria, its connections with basic research, and the research on processes of change, meanwhile the rest of criteria are not met yet. Such a treatment can be considered a strong source of conceptual, empirical and technological innovation, but not of proven efficacy, efficiency and effectiveness yet. It provides concepts and data from basic research that could have a heuristic value for the development of procedures leading to more robust results and hence, with more potential to be widely disseminated.

We could also find a treatment well-established at empirical level regarding the research on processes of change and treatment outcomes (efficacy, effectiveness, and efficiency), but still pending of the conceptual coherence and the establishment of the connections with the basic research. This would be an example of a treatment with robust clinical utility but difficult to disseminate and with a pending potential for innovation.

Likewise, we can have a treatment well established at the level of basic and outcomes research (efficacy, effectiveness, and efficiency) thus easy to disseminate, but whose clinical impact and potential for innovation are to be established yet.

CONCEPTUAL CRITERIA

Table 1 present the conceptual criteria. These represent an innovation in relation to the Chambless Task Force criteria which do not contemplate them. The underlying assumptions of these criteria are:

1. It is not possible to design, implement and do research on a treatment without having an underlying assumption. Every treatment somehow presupposes its target population, the change process involved, its behavioural change goals, and a desirable change. Without specifying these aspects, a treatment cannot be implemented, nor taught, nor tested in a clinical trial.
2. Psychological treatments utilize technical language that may be defined by its precision, scope, organization, and depth (Hayes, 1991). A paradigm for the development of technologies for behavioural change based solely on efficacy data can be defined as a language rich in precision, but poor in the rest of dimensions. This may prevent the treatment from further technological refinements, and strengthen the impact of economic factors on the development of treatments (Rosen & Davison, 2003).
3. As a technical language, every psychological treatment has evolved in a given historic and social context. The key concepts of a treatment (change processes, goals, concept of change), have historic and social foundations that should be made explicit in order to calibrate its actual utility in the context it evolved.

Table 1. Conceptual Criteria for the Validation of Psychological Treatments.

A. Conceptual criteria
The treatment to validate, in its original manual (or the correspondent publication), and through the research of independent research teams, should:
<ol style="list-style-type: none"> 1. Specify its underlying assumptions, its analytic goals, and its truth criterion. 2. Identify its connections with a basic research program or with the psychological processes already available in the literature. 3. Define the concept of psychopathology and its limits, as well as the data from either the clinical or the basic research that allowed its synthesis. 4. Make explicit its historical relations with alternative concepts of psychopathology. 5. Define its goals, as desirable behavioural changes, given the underlying concept of psychopathology.
Describe the social and professional context where the treatment emerged, and how such context influenced the choice of its behavioural change goals.
Well-established treatment
Every treatment meeting all the criteria is well established at conceptual level.
Treatment to be established yet
If only the criteria 1, 2, 3, and 5 are met, then the treatment is to be established yet at conceptual level.

As Table 1 shows, the conceptual criteria require from a treatment the identification of its basic assumptions, its relation with the basic research on psychological processes, the underlying concept of psychopathology along with its history, the selection of its desirable outcomes, and the description of the social context in which such treatment emerged (anthropological concept of suffering, professional context, etc.). Meeting these criteria means the conceptual area of the treatment is well established. Contrarily, if neither the history of the concept of psychopathology nor the social context where the treatment emerged is specified, it is concluded that the conceptual area of the treatment is to be established yet.

Making explicit the history and the social context of a given concept of psychopathology, allows the description of the context in which a treatment evolved and the calibration of its actual utility. As part of the theoretical development of a treatment, it is important to establish the degree of correspondence between the clinical and the research behaviours constituting its procedures, as well as the social and cultural practices in a given moment. This helps to identify some of the sources of bias inherent to the treatment, and to assert whether or not the results in the controlled trials are due to its effectiveness more than to its efficacy. It is important to consider the social and professional variables that have contributed to the current concepts of mental health, the configuration of the professional boundaries among disciplines as well as the work-effort allocation among them, and the selection of procedures for the behavioural change which are socially acceptable (for instance, Abbott, 1988; Goldstein, 2001).

CONNECTIONS BETWEEN THE TREATMENT AND THE BASIC RESEARCH ON PSYCHOLOGICAL PROCESSES

This set of criteria emphasizes the mutual entailment between the basic science and the technological development, and also constitutes an innovation in relation to the Chambless Task Force criteria. A detailed description of our criteria is presented in Table 2. Their underlying basic assumptions are:

1. The psychological processes involved in human suffering are common to all individuals. In other words, there is no difference between the behaviour of the general and the clinical population in terms of their functions, our analytic goal. The definition of such psychological processes is important for the development of a behavioural technology to the extent that they can explain behavioural events in the daily life taking into account variables whose functions are not obvious (i. e., the function of the suppression strategies, the function of searching coherent narratives, etc.) (Lee, 1988).
2. The knowledge from the basic research is applicable, that is, its goal is the synthesis of functional relations between the behaviour of the individuals and the historical and present conditions. In fact, considering the clinical research and practice as the contexts where the basic knowledge is tested in singular and unrepeatable situations allowing the detection of its limits, is more useful than considering them as mere applications. Besides the fact that the basic research on one hand, and the clinical research and practice on the other hand, are universe of diverse empirical generality contextually controlled by specific variables, they indeed interact in the common language (Carrascoso López, 1999). This has been called model of mutual interest among basic science, technology, and application (Hayes, 1998).

The criteria in Table 2 constitute an operational definition of the previous assumptions. In order to conclude that the connections between the basic research and a particular psychological treatment are well established, such treatment has to specify:

- a. The data from the basic research that have allowed the development of working definitions of the presupposed change processes.
- b. The experimental research on the psychological processes involved in the change processes comprising a treatment for its target population.
- c. The descriptive research showing whether the functional structure of the psychological processes relevant in the treatment is similar across clinical and general populations.

Meeting these criteria may allow the maintenance of the connections between the basic science and the clinical practice without any of them losing specificity. This way we can better conclude about the validity of a given technique. For instance, it is usually assumed that thought stopping is indicated in case of repetitive and irrational thoughts causing suffering to the individual. It is considered a relatively efficacious technique (Lozano Oyola, Rubio Zarzuela & Pérez San Gregorio, 1999). In topographical terms, though stopping is usually considered a distraction technique or a positive punishment. However, lab research on thought suppression and its effects by using the

Table 2. Connections with the Basic Research.

B. Connections with the basic research
The psychological treatment to validate should:
1. Specify the data from the basic research (with human and non-human organisms) on psychological processes relevant for its concept of processes of change.
2. Be based on systematic replication studies with group or temporal series designs, carried out by independent labs doing research on psychological processes with the populations target of the treatment. One possible purpose of such research might be the detection of qualitative and quantitative differences between the target and the comparison populations.
Be based on observational and correlational studies carried out by independent labs, which present descriptive analyses of the similarities in the relevant psychological processes across populations.
Well established treatment
If all criteria are met, the connection with the basic research is well established.
Treatment to be established yet
If one or more of the three criteria are not met, then the connection with the basic research is to be established yet.

white bear paradigm developed by Wegner and colleagues (Wegner, 1994) with general and clinical populations, has challenged the utility of the suppression strategies. Wegner showed, for instance, that suppression turns the suppressed thoughts into hyper-accessible contents (Wegner & Erber, 1992), that relaxation while in stress conditions increments the stress levels (Wegner, Broome & Blumberg, 1997), and that the suppressed thoughts may appear in the dreams (Wegner, Wenzlaff & Kozak, 2004). Thus, thought stopping cannot be considered a procedure for the behavioural change of proven efficacy. It was not synthesized from psychological processes, we do not know which change processes are involved in its effects, and the experimental research on such procedure has consistently shown a lack of mid/long term utility, even its adverse effects. It is, thus, a procedure described with some precision, but with little scope, organization and depth, and consequently, whose effects are not well documented.

On the contrary, the adoption of mindfulness strategies as relevant ingredients within the third wave of behaviour therapies (Hayes, 2004) constitutes a good example of the utility of the basic research (in this case on rule-governed behavior and relational framing) as applicable knowledge. The working definitions of such strategies have derived from the basic research on psychological processes (Zvolensky, Feldner, Leen-Feldner & Yartz, 2005). For that reason, the connection between the basic science and the development of treatments constitute one of the main sources of innovation within the empirical clinical psychology (Hayes, 2005).

The basic research we are contemplating here is not only experimental. The descriptive analysis of the people's experiences (regardless whether we are dealing with

general or with clinical population) is a very useful tool that, although does not allow the direct influence over the relevant behavioural processes, it makes possible the precise identification of target behaviours. The direct and cautious observation of relevant behaviours in-situ through procedures like the experience sampling, may help identify the topographical characteristics of a behaviour, and the contextual variables that should be further explored in the experimental context (de Vries, 1992). In fact, the descriptive research can be relevant as a first step for the systematic identification of psychological processes in the daily life.

RESEARCH ON CHANGE PROCESSES

Table 3 presents these criteria, which constitute another important innovation with respect to the Chambless Task Force criteria. Their underlying assumptions are:

1. The daily clinical practice requires from the basic research the identification of the variables susceptible to be manipulated, and the change processes likely involved in the results of a particular treatment (Hayes, Barlow & Nelson-Gray, 1999). In that sense, rather than the research on the treatment efficacy, which specifies the target population and the average result of the treatment implementation in ideal conditions, thus testing the final results of a treatment with a meta-client, clinicians need broader and more flexible information about the rationale of the work to be done with a given client.
2. The research on the change processes presupposed in a treatment contributes to the establishment of the connections between the basic research on psychological processes and the development of treatments. Also, it turns the daily practice into an easier activity. The change processes are understood as the synthesis of the basic knowledge, which allows the detection of the processes involved in the emergence and maintenance of a given psychopathology. The operational definition of such change processes requires the specification of the variables susceptible to be manipulated, and the behaviours of the professionals and the clients relevant on that regard.

This set of criteria is built in an operational way. It specifies the minimum amount of evidence required and the methodological conditions the studies have to meet for a treatment to be considered well established at this level. The inclusion of these criteria is important and necessary when dealing with the development of techniques because of several reasons. First, it involves the specification of the analytic goals of the treatment model from the very beginning, as well as its relation with the basic research on psychological processes. Second, it provides flexible and rigorous definitions of the relevant behaviours of the therapist and the client, and the variables that the treatment can potentially change, turning the treatment into an easier tool to teach and disseminate. Third, the description of a change process has more empirical generality than the description of a specific technique (systematic desensitization, for instance), and thus, it is widely applicable to diverse problems topographically different but sharing the same functional properties. The result is a definition of treatment which is more flexible and better connected to the psychological models of change. In this sense, the consideration of a treatment yet as a change principle yet as a whole treatment

*Table 3. Research on Processes of Change.***C. Research on processes of change**

For a treatment to be validated at this level:

1. There should be at least two group studies (within or between subjects comparisons), with enough statistical power ($n \geq 30$ subjects per group), and carried out by independent labs with clinical populations, for each process of change presupposed.
2. There should be at least 4 studies with a potent temporal series design ($n \geq 10$), carried out by independent labs with clinical populations, for each change process presupposed.
3. In the case of relying on analogue studies, these will be carried out with humans and according to the methodological requirements specified in 1 and 2. As well, the analogue studies will be used as contrast to determine the potential contribution of each of the processes of behavioural change, and to determine whether such processes change are of a continuous nature.

There should be at least two studies on treatment withdrawal employing experimental designs as indicated in 1 and 2. These studies will confirm the involvement of the process of change presupposed, or contrarily, will show evidence in favour of other processes.

Well-established treatment

Every treatment meeting all the criteria is well established at conceptual level.

Treatment to be established yet

If only the criteria 1, 2, 3, and 5 are met, then the treatment is to be established yet at conceptual level.

package, is irrelevant to the extent that the key is the conceptual and operational definition of the change processes involved. Fourth, the emphasis in the research on change processes would improve the treatment manuals because the relevant clinical behaviours of the therapist and the client would be described in a more flexible and subtle fashion.

TREATMENT OUTCOMES: EFFICACY, EFFECTIVENESS AND EFFICIENCY

Tables 4, 5, 6 present these criteria. There are two main differences with respect to the Chambless Task Force criteria regarding outcomes research. First, ours emphasize the research on the treatment efficiency and effectiveness. Second, we propose that the case studies meeting certain composition rules (for instance, Buena Casal & Sierra, 2002) can serve as treatment effectiveness evidence. This is methodologically risky, but in the case of following some specific rules, the case studies can illustrate the treatment effectiveness. Nonetheless, this proposal requires a meta-analysis of the literature which compared, for instance, the effect size of the efficacy and effectiveness of well-controlled

Table 4. Research on Treatment Efficacy.

D. Research on efficacy outcomes
For a treatment to be validated at this level:
1. There should be at least two group studies (within or between subjects comparisons), with enough statistical power ($n \geq 30$ subjects per group), and carried out by independent labs. Such studies have to show better effects than those utilizing medication, psychological placebo, or any other treatment, or at least, results equivalent to those obtained by other well established treatments.
2. There should be at least 4 studies with a potent temporal series design ($n \geq 10$), carried out by independent labs. Such studies have to show better effects than those utilizing medication, psychological placebo, or any other treatment, or at least, results equivalent to those obtained by other well established treatments.
Both the group studies and the temporal series studies have to use a treatment guide, to control the internal validity, and to specify the characteristics of the clients assigned to the experimental and the control conditions.
Well established treatment
If all criteria are met, then the treatment efficacy is well established.
Treatment to be established yet
A treatment is likely to prove its efficacy if: a) the studies meet all criteria but have been carried out by the same research team; b) the available studies employ the waiting list as the control condition; c) the studies do not meet the qualitative and quantitative standards required in 1, 2 and 3.

studies, to the effect size of the case studies. The basic assumptions of these criteria are:

1. The research on efficacy provides preliminary data on the effect of a treatment in conditions which are standard and replicable across independent investigators. In this sense, it allows a restricted testing of the change principles at the heart of the design of a specific treatment.
2. The data from the research on the treatment effectiveness define the limits of its efficacy, facilitating the identification of the conditions in which such treatment is more or less efficacious, and taking into account variables related to the clients as well as to the professionals. In other words, it provides information about the generality of the change principles assumed by a treatment.
3. The research on treatment efficiency contributes to the identification of the economical, social and cultural relevance of a treatment, for the clients, the professionals, and the politicians. That is, it allows the research on the social relevance of the change principles comprising the treatment.

Probably, the research on treatment effectiveness and efficiency has more impact for the governmental decision-making regarding sanitary issues, than research on treatment efficacy. Research on efficacy informs about which psychological treatments are better as an average. But it does not inform about the extent to which the conditions for the application of the treatment influence its effects, whether the treatment is easy or difficult to train and disseminate among clinicians and researchers, and what might be

Table 5. Research on Treatment Effectiveness.

E. Research on effectiveness outcomes

For a treatment to be validated at this level:

1. There should be at least 2 group studies (within or between subjects comparisons), with enough statistical power ($n \geq 30$ subjects per group), and carried out by independent labs. The studies should show that the effects of the treatment in the efficacy trials are, at least, equivalent both statistically and clinically. Such studies should be carried out with different populations, by different therapists, and in real professional contexts.
2. There should be at least 4 studies with a potent temporal series design ($n \geq 10$), carried out by independent research teams. Results should show that the effects of the treatment in the efficacy trials are, at least, equivalent both statistically and clinically. Such studies should be carried out with different populations, by different therapists, and in real professional contexts.

There should be, at least, 30 case studies written according to normative guidelines, and published by different therapists working in different institutions and socio-cultural environments.

Well-established treatment
If all criteria are met, then the treatment effectiveness is well established

Treatment to be established yet
The treatment effectiveness is to be proven yet if: a) the studies meet all criteria but have been carried out by the same research or clinical team; b) the available studies employ the waiting list as the control condition; c) the studies do not meet the qualitative and quantitative standards required in 1, 2 and 3.

Table 6. Research on Treatment Efficiency.

F. Research on efficiency outcomes

For a treatment to be validated at this level:

1. There should be at least one well-controlled study on efficacy and one well-controlled study on effectiveness, where the costs in the short, mid and long run are calculated. As well, such costs should be calculated for the general population.
2. There should be at least two studies aimed at identifying the variables promoting the acceptance or the rejection of the well-established treatments on the side of the clients.

Well-established treatment
If the two criteria are met, then the treatment efficiency is well-established.

Treatment to be established yet
The treatment efficiency is to be established yet if: a) the studies meet all criteria but have been carried out by the same research or clinical team; b) the available studies employ the waiting list as the control condition; c) the studies do not meet the qualitative and quantitative standards required in 1 and 2.

the effect of training and disseminating in terms of improvement of the daily clinical practice. A treatment that is easy training and disseminating, with an effect size not dependable of the conditions of implementation, and with potential to produce important changes in the performance of sanitary professionals, can turn into a very attractive tool for professionals and politicians. Studies by Hayes, Bissett, et al. (2004), Strosahl, Hayes, Bergan, and Romano (1998), and the reviews by Cummings, O'Donohue, and Ferguson (2002), constitute a good example on this regard.

Research on treatment efficiency should not be focused exclusively in the research on economic costs. "Treatment cost" can be operationally defined from diverse points of view. To the sanitary administrators, research on economic costs can be very useful. In fact, studies on efficiency have been usually focused on the economic impact of the treatment. However, treatment cost can also be defined in terms of the client's implication. For instance, a question still open is whether a treatment requiring high implication on the side of the client would be more efficient than a treatment perfectly fitting the cultural practices and goals that the client brings to therapy. This way, the research on the treatment efficiency is related to the sociologic and anthropologic dimensions of the clinical practice. The data analyzed by Seligman (1995) are a good indicator of this: clients are happy regardless of the treatment implemented. Beyond criticisms, psychoanalysis has adepts still. As Pérez Álvarez (1992) showed, it constitutes a socio-cultural institution shared by clients and professionals, in this sense it could be a good example of an efficient treatment. On the other hand, and regardless of its efficacy, cognitive-behavioural therapy is not the standard treatment offered to clients yet. Regarding the studies on efficiency, then, we may question to what extent a treatment establishes contact with the social and cultural practices in a given historic moment; and if that was the case, whether this continuity has any impact on the clients quality of life, and in which direction. We are explicitly considering that the research on the treatment efficiency is a multi-disciplinary field, whose main function is to empirically calibrate the relevance of the data on the treatment efficacy and effectiveness.

Given the functional nature of the present alternative criteria, we claim that research on treatment effectiveness and efficacy should consider the inclusion of improvement indicators different to symptoms reduction, as it has been mentioned in the introductory section of the paper. The big enterprise then, is to define what would be a desirable pattern of change, and develop the assessment strategies that would better capture such a pattern. Having into account individual criteria more than group criteria, research on treatment effectiveness and efficacy might move from symptoms management to more functional criteria like behaving in accordance to personally valued directions even in the presence of thoughts and feelings experienced as aversive (Luciano, 2001), achievement of personal goals, and the like. This type of patterns of changes seems to contribute to deep and stable improvements in the quality of life of the individuals, understanding this as being responsible for the choices the client makes in his/her daily living. This is a dependent variable interesting for several therapy traditions, from the existentialism to the third wave of behaviour therapies, and there are several efforts to define it in operational terms as well as to measure it in a reliable way (see review in Hayes *et al.*, 2004; Luciano, 2001).

CONCLUSION

We have presented a set of criteria for the validation of psychological treatments alternative to those by the Chambless Task Force, standard criteria used in the research on treatment outcomes. Our criteria are explicitly based on the assumption that the development of a psychological treatment is a very complex process encouraged by the assumptions of clinicians and researchers, that has to be connected to the theory and the basic research, and open to the diversity of individual change patterns, out of which the reduction of frequency, intensity, and duration of a limited number of symptoms is only one of the possibilities. The claimed incorporation of improvement indicators different to symptoms reduction, according to the functional perspective underlying the present criteria, deserves two considerations. First, it implies the incorporation of new ways of defining and measuring the expected patterns of change (as in Luciano, 2001); second, selecting certain improvement indicators over others responds to conceptual and theoretical assumptions regarding psychopathology and therapy (Paéz-Blarrina *et al.*, 2006), as this should not be forgotten neither left implicit. Contrarily, specifying its underlying assumptions will facilitate the training, implementation, evaluation, and dissemination of the psychological treatment at play. In relation to the evaluation of the treatment's outcomes, we understand it is a complex and multidimensional process that should allow the selection of clinical strategies and of the conceptual approaches that facilitate the incorporation of the dependent variables best revealing the changes in the life of the clients.

Likewise, we think it is important to turn the definition of treatment into a more flexible and less medical concept. A treatment may be defined either as the systematic and strategic implementation of a change principle, or as a whole set of techniques. Either way, the key is the specification of the relevant change processes, and the operational definition of the clinical relevant behaviours of professionals and clients, not simply the number of techniques included in such treatment, or its formal specification.

Finishing up, we highlight that the present proposal is still open to further considerations. As mentioned before, psychopathology according to the traditional diagnostic criteria has to do with diverse forms of behaving that produce discomfort or interference across life domains. We may wonder then, whether the actual symptoms of mental illness would not be consider us such in the case the interference disappeared. As well, we will have to take into account the role of psychosocial contingencies in shaping the psychodiagnostic categories, as the historical, epidemiological and sociologic research has pointed out. Adopting a functional perspective will allow developing functional psychopathology criteria and a truly psychological clinical psychology (Pérez Álvarez, 2003), as well as departing from the medical criteria widely utilized at present. Such is the direction the present alternative criteria points to.

REFERENCES

- Abbott A (1988). *The system of professions. An essay on the division of expert labour*. Chicago: The University of Chicago Press.
- Berrios GE (1984). Descriptive psychopathology: Conceptual and historical aspects. *Psychological Medicine*, 14, 303-313.
- Berrios GE & Porter R (Eds.) (1995). *A history of clinical psychiatry*. New York: Athlone Press.
- Bourgeois M (1994). The first few minutes: Original contact and the speed of psychiatric diagnosis. In P Pichot & W Rein (Eds.), *The Clinical Approach in Psychiatry* (pp. 273-286). Paris: Les Empêcheurs de Penser en Rond.
- Buela Casal G & Sierra JC (2002). Normas para la redacción de casos clínicos. *International Journal of Clinical and Health Psychology*, 2, 525-532.
- Callahan CM & Berrios GE (2005). *Reinventing depression. A history of the treatment of depression in primary care (1940-2004)*. New York: Oxford University Press.
- Carrascoso López FJ (1999). A conceptual analysis of the problem of the relationship between basic research and applied work in psychology: The examples of experimental and applied behavior analysis. *Psychology in Spain*, 3, 3-24.
- Chambless DL, Baker MJ, Baucom DH, Beutler LE, Calhoun KS, Crits-Cristoph P et al. (1998). Update on empirically validated therapies, II. *The Clinical Psychologist*, 51, 3-16.
- Chambless DL & Hollon SD (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66, 7-18.
- Chambless DL & Ollendick TH (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52, 685-716.
- Chambless DL, Sanderson WC, Shoham V, Bennet-Johnson S, Pope KS, Crits-Cristoph P et al. (1996). An update on empirically validated therapies. *The Clinical Psychologist*, 49, 5-18.
- Cummings NA, O'Donohue WT & Ferguson KE (2002). *The impact of medical cost offset on practice and research: Making it work for you*. Reno: Context Press.
- de Vries M (1992). *The experience of psychopathology. Investigating mental disorders in their natural settings*. Cambridge: Cambridge University Press.
- Follette WC, Bach PA & Follette VM (1993). A behaviour-analytic view of psychological health. *The Behavior Analyst*, 16, 303-316.
- Goldfried MR & Wolfe BE (1998). Toward a more clinically valid approach to therapy research. *Journal of Consulting and Clinical Psychology*, 66, 143-150.
- Goldstein J (2001). *Console and classify. The French psychiatric profession in the nineteenth century* (second edition). Chicago: The Chicago University Press.
- Hayes SC (1991). The limits of technological talk. *Journal of Applied Behavior Analysis*, 24, 417-420.
- Hayes SC (1998). Building a useful relationship between "applied" and "basic" science in behaviour therapy. *The Behavior Therapist*, 21, 109-112.
- Hayes SC (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy*, 35, 639-665.
- Hayes SC (2005). Eleven rules for a most successful clinical psychology. *Journal of Clinical Psychology*, 61, 1055-1060.
- Hayes SC, Barlow DH & Nelson-Gray R (1999). *The scientist practitioner. Research and accountability in the age of manager care* (second edition). Boston: Allyn and Bacon.
- Hayes SC, Bissett R, Roget N, Padilla R, Kohlenberg BS, Fisher G et al. (2004). The impact of acceptance and commitment training and multicultural training on the stigmatizing attitudes and professional burnout of substance abuse counselors. *Behavior Therapy*, 35, 821-835.

- Hayes SC, Strosahl, K. D., Bunting, K., Twohig, M., & Wilson, K.G. (2004). What is Acceptance and Commitment Therapy? In SC Hayes & KD Strosahl (Eds.), *A practical guide to Acceptance and Commitment Therapy* (pp. 1-30). New York: Springer.
- Lee VL (1988). *Beyond behaviorism*. Hillsdale: LEA.
- Lozano Oyola JF, Rubio Zarzuela EM & Pérez San Gregorio MA (1999). Eficacia de la técnica de detención del pensamiento en diferentes trastornos psicopatológicos. *Psicología Conductual*, 7, 471-499.
- Luciano MC (Ed.) (2001). *Terapia de Aceptación y Compromiso: Libro de casos*. Valencia: Promolibro
- Páez Blarrina M, Gutiérrez Martínez O, Valdivia Salas S & Luciano C (2006). ACT y la importancia de los valores personales en el contexto de la terapia psicológica. *International Journal of Psychology and Psychological Therapy*, 6, 1-20.
- Pérez Álvarez M (1992). *Ciudad, individuo y psicología. Freud detective privado*. Madrid: Siglo XXI.
- Pérez Álvarez M (2003). *Las cuatro causas de los trastornos psicológicos*. Madrid: Universitas.
- Rosen GM & Davison GC (2003). Psychology should list empirically supported principles of change (ESPs) and not credential trademarked therapies or other treatment packages. *Behavior Modification*, 27, 300-312.
- Seligman MP (1995). The effectiveness of psychotherapy. The consumer reports study. *American Psychologist*, 50, 965-974.
- Strosahl KD, Hayes SC, Bergan J & Romano P (1998). Assessing the field effectiveness of Acceptance and Commitment Therapy: An example of the manipulated training research method. *Behavior Therapy*, 29, 35-64.
- Task Force on Promotion and Dissemination of Psychological Procedures (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, 48, 3-23.
- Wegner DM (1994). *White bears and other unwanted thoughts. Suppression, obsession, and the psychology of mental control* (second edition). New York: Guilford Press.
- Wegner DM, Broome A & Blumberg SJ (1997). Ironic effects of trying to relax under stress. *Behaviour Research and Therapy*, 35, 1, 11-21.
- Wegner DM & Erber R (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology*, 63, 903-912.
- Wegner DM, Wenzlaff RM & Kozak M (2004). Dream rebound: The return of suppressed thoughts in dreams. *Psychological Science*, 15, 232-236.
- Westen D, Novotny CM & Thompson-Brenner H (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631-663.
- Wilson KG & Luciano MC (2002). *Terapia de aceptación y compromiso. Un tratamiento conductual orientado a los valores*. Madrid: Pirámide.
- Zvolensky MJ, Feldner MT, Leen-Feldner EW & Yartz AR (2005). Exploring basic processes underlying Acceptance and Mindfulness. In SM Orsillo & L Roemer (Eds.), *Acceptance and mindfulness-based approaches to anxiety. Conceptualization and treatment* (pp. 325-358). New York: Springer.

Received, September 1 2007

Final Acceptance, September 30 2007