

DEVELOPMENT OF SHORT SCREENING INSTRUMENTS FOR PSYCHIATRY, PSYCHOTHERAPY, AND CRISIS INTERVENTION

Andreas Maercker & Hans Menning

Empirical investigation of the effectiveness and efficacy of different forms of therapy allow more informed treatment decisions to be taken, particularly as pharmacotherapy research sets very high standards for the demonstration of efficacy (Lafata et al. 1996). In the first section of the present chapter, some central observations on the measurement of therapeutic outcome will be presented. In the second section, a core battery of instruments (called „Bel-Air“) developed for the Department of Psychiatry at the University of Geneva for the evaluation of psychotherapy will be described. Finally, in the third section, screening instruments for the quick assessment of post-traumatic stress disorder (PTSD) will be presented.

THERAPY RESEARCH: CENTRAL ISSUES AND FINDINGS

Despite the wealth of research findings, there is still no clear evidence on the comparative efficacy of different therapies (Guimón 1997). Potentially useful ways of determining the differential effectiveness of treatment methods are randomized controlled trials or meta-analysis, the latter a procedure in which the effect sizes of different therapies

are compared. Meta-analysis allows empirical data from a large set of individual investigations to be summarized, and quantitative conclusions to be drawn. The main precondition is the availability of controlled studies in which a treatment condition is compared with a control condition (e.g., occupational therapy groups, regular psychiatric wards).

Concerning psychiatric group therapy, a meta-analytic review by McRoberts, Burlingame, and Hoag (1998) encompassed 23 studies which directly compared the effectiveness of the individual and group therapy formats. Like the majority of previous meta-analytic reviews, the results of this analysis indicated that there is little difference in efficacy between individual and group therapy.

The confounding of effect sizes with researcher allegiance is a controversial finding in meta-analyses (Lambert & Bergin 1994) and underscores the need for rigorous methodology to minimize bias in outcome research. Thus, assessment methods should be applicable to different research centers or hospitals in various locations (i.e., multi-site studies) to control for allegiance bias.

Initiatives have recently been taken to enhance the methodology and design of research into the efficacy of psychological therapy. These include the introduction of standardized systems to document patient characteristics (e.g., Kallert et al. 2000) and attempts to standardize psychiatric instruments measuring therapeutic outcomes (Piersma & Boes 1997). Recently, standardized „real-time monitoring“ systems (e.g. the „Synergetic Navigation System“, SNS) for detailed patient documentation and process-outcome-measurements via handhelds or mobile phones has been introduced (Schiepek et al. 2005). This kind of internet-based system enables a fast, broad and frequent assessment of the therapy process and outcome. However, this system is currently yet under evaluation.

In the past, the diversity and variability of the classification systems reflected its non-binding nature. A theoretical foundation had not been made. To date, as regards design of data collection and the definition of criteria, there is an increased need of standardization of

procedures, at least in research. The use of validated assessments at several points in time (at least pre and post treatment) and follow-up measurements are seen as indispensable. As a result of methodological developments of success criteria, it is possible to reach general agreement on several indices of treatment success: effect size, reliable change, and clinical significance.

Three fundamental assumptions constitute the basis of the methodological distinction of outcome levels: (i) the prescriptive decision for the treatment goal of healing or improvement of disorders (versus other meta-theories of altered psychological states, e.g., labeling theories); (ii) on the basis of this decision the choice of a generic model of diseases as the common basis determination of the content of outcome variables and the decision for appropriate methodology.

There is a general consensus that outcome assessment should be based on multiple criteria; it should be multidimensional, multimethodal and multimodal. Moreover, the following steps should be taken in constructing a research framework: In a first step, the theoretical background should be considered in order to identify the conceptual variables or indicators to be used in the assessment of therapeutic outcome. In a second step, ways of operationalizing and measuring these variables should be determined. This differentiation between conceptual and methodological aspects is of great importance in the definition of therapy outcomes, since each aspect is independent of the other.

2. OBJECTIVES OF THE “BEL-AIR” CORE BATTERY

Which conceptual *dimensions* are appropriate for evaluating the success of therapeutic outcomes? Usually, the outcome of therapy is measured on the following three levels:

- 1) etiology of dysfunction or deficiencies
- 2) symptomatology — both disorder-specific and general
- 3) social roles such as the "sick-role" or "impairment of the normal role".

Up to now, the *etiology* of dysfunction or deficiencies have been largely overlooked in evaluations of the outcomes of therapeutic interventions (e.g., psychotherapy, pharmacotherapy). According to current knowledge on psychiatric disorders, genetic, biochemical, physiological, psychological, or sociological factors may operate here. As yet, little use has been made of the opportunity to assess biological and neurophysiological parameters by means of neuro-imaging techniques such as PET and (f)MRI in this context (but cf. Krishnan 1999). In contrast, some of the psychological factors involved in the maintenance of the disorder (e.g., expressed emotions, cognitive dysfunctional attitudes) are assessed in therapy studies (e.g., NIMH Treatment of depression study: Shea et al. 1992).

Symptomatology can be tapped directly using symptom inventories such as the DSM (American Psychiatric Association, 1994) and the ICD-10 (World Health Organization, 1992), as well as validated observer-rating and self-report instruments screening for symptoms of psychopathology (e.g., SCID, SCL-90-R).

The level of *social roles* embraces aspects of the quality of life as well as interpersonal or social interaction. In recent years, measurements of the quality of life, in particular, have become established. As an additional dimension, the *resources* and coping skills acquired or “rediscovered” during the therapy might be considered. There is now general consensus that the evaluation of treatment outcomes should distinguish between at least these three conceptual levels, and that the causes of dysfunction have thus far been neglected.

In the past decades, repeated attempts have been made to identify appropriate measures for each of the individual levels and thus to construct a standardized inventory for group evaluation. The core battery developed by the American Group Psychotherapy Association is one example (Dies & MacKenzie 1983). This standardized inventory consists of a symptom checklist (SCL-90-R), a scale to measure social situation and emotional problems (Emotions Profile Index), a scale for the global assessment of psychosocial impairment (Global Assessment Scale and General Improvement Rating), and a rating of the patients' individual problems during the therapy process. This core battery did not meet with broad acceptance, however, as the interests of individual research groups were too diverse (Bednar & Kaul, 1994).

On the basis of these observations and the current debate in psychotherapy research, the following recommendations for the sound assessment of therapy outcomes have been formulated (Strauß et al. 1996):

- The methodology applied should permit the outcomes of different forms of treatment (psychotherapeutic and pharmacotherapeutic) to be compared.
- Evaluation research should avoid infringing upon clinical treatment.
- Patient characteristics should be documented as precisely as possible using standardized procedures (e.g., established systems of basis documentation and diagnostic classification).
- Established (i.e., validated) instruments tapping various levels of potential change (causes and deficiencies, symptoms and social roles) should be used.
- Negative effects should also be recorded in detail (e.g., dropout characteristics).

3. THE 'BEL-AIR' CORE BATTERY FOR THE ASSESSMENT OF PSYCHIATRIC DISORDERS

Today's psychiatric departments involve a variety of treatment settings for psychosocial interventions, e.g., inpatient units, outpatient clinics, day clinics, dispensary departments for patients with chronic symptoms. For the evaluation of therapeutic outcomes, a common set of measurement tools is needed, which can be deployed in each of these settings and with various groups of psychiatric patients. Short screening batteries are a valid, powerful, and time-effective tool to fulfill this demand.

In order to evaluate the effectiveness of different therapies developed in the Department of Psychiatry at the University of Geneva, the outcomes of different treatment methods were assessed on several levels (see section 2): in a global assessment of therapeutic outcomes by patients and therapists and a more differentiated assessment of changes (pre-post treatment) in patient responses to questionnaires screening symptoms and aspects of social roles. An overview of instruments is given in Table 1.

Table 1: Conceptual levels and methodological implementation of an evaluation program for different psychiatric disorders (University of Geneva core battery)

<u>Level of assessment</u>	<u>Variables</u>	<u>Instruments</u>
Symptoms	- Complaints/psychopathology (self-report) - Diagnosis (clinician ratings)	Brief Symptom Inventory (BSI, Derogatis & Melisaratos 1983; BSI, Derogatis 1993) DSM-IV (American Psychiatric Association, 1994)
Social roles	- Personal relationships (self-report) - Social functioning (clinician ratings) - Social functioning (self-report)	Evaluation of Relationships Questionnaire (ERA, Fredenrich & Zinetti, 2000) Global Assessment of Functioning (DSM-IV-Axis V, American Psychiatric Association, 1994) Social Functions Questionnaire (QFS, Zanello et al. 2006)

In the following, the central elements of the core battery — BSI, GAF, ERA, and QFS — will be presented, and reasons for the selection of these particular instruments given.

The Brief Symptom Inventory (BSI, Derogatis & Melisaratos, 1983; Derogatis, 1993) is a brief self-report measure which is appropriate for the assessment of psychopathology and psychological complaints. It is a short form of the SCL-90-R, which was devised by the same authors, and includes only 53 of the original 90 items. Administration time was thus cut from around 15–20 minutes to 5–10 minutes. The items are summarized to form nine symptom scales (somatization, obsessive-compulsive, depression, anxiety, anger-hostility, phobic anxiety, paranoid ideation, interpersonal sensitivity, and psychoticism) and three global indices (global severity index, positive symptom distress index, positive symptom total).

Up to now, the BSI has mainly been implemented in studies involving patients with psychological disorders (for an overview, see Piersma et al. 1994) and in assessments of

treatment outcomes in psycho- and pharmacotherapy (for an overview, see Franke, 1997). The BSI has now been translated into several languages (e.g., French, German, Hebrew, Italian, and Spanish).

The Global Assessment of Functioning (GAF) was selected to assess social roles and social functioning in global terms, using a single measure. GAF equals Axis V of the DSM-IV (American Psychiatric Association, 1994). It was developed and operationalized by Endicott et al. (1976) as global sickness rating scale. The GAF scale is to be rated with respect only to psychological, social, and occupational functioning. The instructions exclude impairment of social functioning due to physical (or environmental) limitations. The GAF considers these functioning on a hypothetical continuum of mental health-illness. The GAF scale may be rated for several time periods (e.g., current or the highest level of functioning for at least a few month during the past year).

The Evaluation of Relationships Questionnaire (French: ERA, Fredenrich & Zinetti, 2000) is a new self-report measure intended to be equally appropriate for different psychiatric and psychotherapeutic disorders (e.g., schizophrenia, depressive and anxiety disorders), and to be shorter and more economical than the existing instruments. The ERA consists of 16 items capturing 4 dimensions (= factors in the factor analysis): 1.) Openness to others, 2.) Relational Distance, 3.) Personal/relational insight, and 4.) Pleasure in/through relationship (Zinetti-Bertschy et al. 2004). The ERA shows a high convergent validity, a systematic relation with clinical features (with BSI $r=0.6$ and QFS $r=0.5$), a high change sensitivity and a nice usability through its shortness and fast practicability

The *Social Functions Questionnaire*, (French QFS, Zanello et al. 2006) was as a short, economical instrument to tap social role functions in different psychiatric disorders. The Social Adjustment Scale (SAS: Weissman & Bothwell 1976) was not selected because it was found too long and too redundant in a pilot study. The QFS captures a total of eight domains, two questions being asked for each (1. frequency of activity, 2. satisfaction with activity): 1.) activities, 2.) everyday tasks, 3.) leisure, 4.) family and pair relationship, 5.) extra-familial relationships, 6.) administrative and financial management, 7.) general health, and 8.) social life and information. There is one validation study which found satisfactory psychometric

properties for the QFS. It possesses equivalent or better qualities than comparable validated instruments (like SASS, Social Adaptation Scale Self Evaluation Scale or SAS-SR (convergent validity ($r=0.65$), the Social Adjustment Self-Report). It seems to be a valuable instrument for the monitoring of social functioning in psychiatric patients because clinical norms are available, it reveals a high change sensitivity (studied for outpatients) and it is shorter, faster, with higher reliability than previous instruments.

All scales outlined here constitute a basic core of outcome measures for various psychiatric populations. During the last decade, disorder specific approaches gained more attention. In the remainder of the chapter, we would like to highlight the area of post-traumatic stress disorder (PTSD) and related conditions which is the core of our own clinical and research expertise (e.g., Maercker, Zöllner, Menning, Rabe, Karl, 2006; Knaevelsrud & Maercker, 2007). Not only PTSD patients are of need of appropriate clinical assessment of trauma aftermath or PTSD symptomatology, but also a variety of other patients with further disorders (e.g., Affective disorders, anxiety disorders, Borderline personality disorder).

4. ASSESSMENT OF ACUTE AND POSTTRAUMATIC STRESS DISORDER IN CRISIS INTERVENTION

Post-traumatic stress symptomatology is handled in its first month as Acute Stress Disorder (ASD), fulfilling the A-criterion of DSM-IV (the traumatic event), embracing most of the PTSD symptoms and, additionally, the dissociation criterion. However, the main distinction of a PTSD is the persistence of symptoms like intrusions, hyperarousal, avoidance, numbing for longer than one month.

But not only the distinction between PTSD and ASD is relevant in the clinical practice, but also possible side branches of posttraumatic developments like Adjustment Disorder (AD), Complicated Grief Disorder (CGD, Forstmeier and Maercker, 2007, Maercker, Einsle and Koellner, 2007), Complex PTSD or even Borderline Personality

Disorder (BPD). For these assessments, two short instruments are presented: PTSS-10 and the Breslau PTSD Screening Scale.

The Posttraumatic Symptom Scale, PTSS-10 is a short screening scale for the assessment of posttraumatic stress consisting of 10 items. The questionnaire is conceived in accordance with the concept of the diagnostic and statistical manual DSM including hyperarousal symptoms. It has originally been developed in Norway after a disaster at an oil drilling rig (Weisaeth, 1989). It is available in 8 different languages and is appreciated for its short duration (less than 3 minutes). It is especially useful for survivors of severe disasters, catastrophes and torture. Holen (1990) demonstrated for the Norwegian version an unifactorial solution.

La Escala de Estrés Postraumático (*The Posttraumatic Symptoms Scale PTSS-10*) es un instrumento de 10 items concebido de acuerdo con el DSM. Fue desarrollado inicialmente en Noruega después de una catástrofe en una plataforma petrolera (Weisaeth, 1989). Está disponible en 8 idiomas diferentes y es muy apreciado porque se puede aplicar en menos de tres minutos. Resulta especialmente útil para los supervivientes de catástrofes graves y de torturas. Holen (1990) utilizó una versión unifactorial para la población noruega.

Sin embargo su sensibilidad es bastante baja (75%) para un umbral de 35, debido a una excesiva inclusión de síntomas depresivos. En contraste, su especificidad es muy alta (98%), así como su valor predictivo positivo y negativo (91 y 93%). La consistencia interna (Alpha Conbachs) y su fiabilidad test-retest son bastante buenas (0.85 y 0.89, Holen, 1990). El PST-10 es ampliamente utilizado actualmente en las Unidades de Urgencia y crisis médicas y psiquiátricas para despistar el estrés postraumático.

However, assumed an optimal threshold value of 35, its sensitivity is rather low (77%), caused by an overinclusion of depressive symptoms. In contrast, the specificity of the test is very high (98%), similarly the positive and negative predictive value (91 and 93%). Internal consistency (Cronbachs Alpha) and test-retest-reliability were rather good (0.85 and 0.89, Holen, 1990). The PTSS-10 is now widely used in Intensive Care or Critical Care Medicine as assessment or screening for posttraumatic stress.

The Breslau PTSD Screening Scale (Breslau et al., 1999) is newer and even shorter assessment or screening means. It consists of only 7 items chosen of the most sensitive and specific DSM-IV criteria. Its sensitivity and specificity is rather good (80% and 97%, respectively), similarly the positive and negative predictive value (71 and 98%). In its 7-item form it does not include intrusive symptoms because of their low specificity. Therefore, for clinical purposes we added 2 additional intrusion items such that all PTSD symptoms are covered. This version is an efficient method to screen for PTSD in research and clinical studies. The Breslau PTSD Screening scale has been widely used in primary care medicine in the US recently.

5. OUTLOOK

As shown in the overview above, great progress has been made in assessment methodology. In particular, the development of short, economical instruments has been a milestone of development in recent years; administration is easier for both patient and evaluator, and they are much better suited to day-to-day clinical practice.

Methods of evaluation will continue to evolve over the years to come. It is to be expected that, even in non-biological (non-pharmacological) therapies, biopsychiatric methods will be used to evaluate causes of dysfunction. Such biopsychiatric methods include parameters of gene expression, neurotransmitters and psychophysiology (e.g., cerebral blood flow; Crespo-Facorro et al. 1999) Within the field of psychiatry, however, the psychological causes of dysfunction and deficiency will continue to play a pivotal role (e.g., expressed emotions parameters, cognitive misinterpretations of basic symptoms, dysfunctional interpersonal problem solutions).

Another possible conceptual level of assessment (cf. Table 1) is the social level of shared cultural values (Guimon et al. 1999; Maercker & Müller 2000). This aspect is of particular relevance with respect to the question of whether evidence-based therapeutic

methods can be used to treat psychological disorders in countries with more limited financial resources.

6. REFERENCES

- Crespo-Facorro B, Cabranes J, Lopez-Ibor M, Paya B, Fernandez Perez C, Encinas M, Ayuso Matoes J, Lopez-Ibor JJ (1999) Regional blood flow in obsessive-compulsive patients with and without a chronic tic disorder. A SPECT study. *European Archives of Psychiatry and Clinical Neuroscience* 249: 156-161
- Derogatis L (1993) Brief Symptom Inventory (BSI), administration, scoring, and procedures manual. National Computer Services, Minneapolis
- Derogatis L, Melisaratos N (1983) The Brief Symptom Inventory: An introductory report. *Psychological Medicine* 13: 595-605
- Dies R, MacKenzie K (1983) *Advances in group psychotherapy*. International Universities Press, New York
- Endicott J, Spitzer R, Fleiss J, Cohen J (1976) The Sickness Rating Scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry* 33: 766-771
- Forstmeier S & Maercker A (2007) Comparison of two diagnostic systems for complicated grief. *J Affective Disorders* 99: 203-211.
- Guimón J (1997) *Thérapie groupale dynamique intensive et brève*. *Psychothérapies* 18: 15-21
- Guimon J, Fischer W, Sartorius N (1999) *The images of madness. The public facing mental illness and psychiatric treatment*. Karger, Basel
- Kallert T, Schutzwahl M, Leisse M, Kluge H, Kilian R, Angermayer M, Bach O (2000) [Standardized documentation system for the complementary sector of psychiatric care. Development and trial in Saxony]. *Psychiatrische Praxis* 27: 86-91
- Knaevelsrud C & Maercker A (2007) Internet-based treatment for PTSD reduces distress: a randomized controlled clinical trial. *BMC Psychiatry*, 7,13:1-10.
- Krishnan K (1999) Brain imaging correlates. *Journal of Clinical Psychiatry* 60 (Suppl.): 50-54
- Lafata J, Koch G, Ward R (1996) Synthesizing evidence from multiple studies. The role of meta-analysis in pharmacoeconomics. *Medical Care* 34 (Suppl.): 136-145
- Lambert M, Bergin A (1994) The effectiveness of psychotherapy. In: Bergin A, Garfield S (eds) *Handbook of psychotherapy and behavior change*. Wiley, New York, pp 143-189
- Maercker A, Müller J (2000) Recent developments in cognitive-behavioral treatment of posttraumatic stress disorder. *Italian Journal of Psychiatry and Behavioral Science*
- Maercker A, Zöllner T, Menning H, Rabe S, Karl A (2006) Dresden PTSD treatment study: randomized controlled trial of motor vehicle accident survivors. *BMC Psychiatry*, 6, 29:1-8.

Maercker & Menning 12

- Maercker A, Einsle F, Köllner V (2007) Adjustment disorders as stress response syndromes: A new diagnostic concept and its exploration in a medical sample. *Psychopathology* 40: 135-146.
- McRoberts C, Bulingame G, Hoag M (1998) Comparative efficacy of individual and group psychotherapy: A meta-analytic perspective. *Group Dynamics* 2: 101-117
- Piersma H, Boes J (1997) MCMI-III as a treatment outcome measure for psychiatric inpatients. *Journal of Clinical Psychology* 53: 825-831
- Schiepek G, Eckert H, Weihrauch S (2005) Prozessmonitoring dynamischer Systeme. *Systeme* 19: 177-209
- Shea M, Elkin I, Imber S, Sotsky S, Watkins J, Collins J, al. e (1992) Course of depressive symptoms over follow-up. Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Archive of General Psychiatry* 49: 782-787
- Strauß B, Eckert J, Tschuschke V (1996) *Methoden der empirischen Gruppentherapieforschung [Methods of empirical group therapy research: A handbook]*. Westdeutscher Verlag, Opladen
- Weissman M, Bothwell S (1976) Assessment of social adjustment by patients self-report. *Archives of General Psychiatry* 33: 1111-1115
- Zanello A, Weber Rouget B, Gex-Fabry M, Maercker A, Guimon J (2006) Validation du Questionnaire de fonctionnement social (QFS), un autoquestionnaire mesurant la fréquence et la satisfaction des comportements sociaux d'une population adulte psychiatrique. *L'Encéphale* 32: 45-59
- Zinetti-Bertschy A, Fredenrich-Muhlebach A, Rosner M, Guimon J, Maercker A (2004) First steps towards the validation of the Relations to Others French Scale (ERA: Echelle de Relations avec Autres). *European Journal of Psychiatry* 18 (Suppl.): 85-92

Traducción

DESAROLLO DE INSTRUMENTOS BREVES DE DESPISTAJE PARA LA PSIQUIATRIA, LA PSICOTERAPIA Y LA INTERVENCIÓN EN CRISIS

Andreas Maercker & Hans Menning

La investigación empírica sobre la eficacia y la efectividad de diferentes formas de terapia permite que se tomen decisiones más informadas acerca del tratamiento debido, particularmente, a que la investigación farmacoterapéutica establece unos niveles de exigencia muy elevados para la demostración de eficacia (Lafata et al. 1996). En la primera sección del

presente capítulo, se expondrán algunas observaciones centrales sobre la medición de los resultados terapéuticos. En la segunda sección, se describirán una batería de tests para la evaluación psicoterapéutica (llamada *Bel-Air*), elaborada para el Departamento de Psiquiatría de la Universidad de Ginebra. Y finalmente, en la tercera parte, se presentarán las pruebas de despistaje de valoración rápida del Trastorno por Estrés Postraumático (TEPT).

1. Investigación terapéutica: cuestiones centrales y conclusiones

A pesar de la abundancia de hallazgos en el campo de la investigación, todavía no existe evidencia clara sobre la eficacia comparada de las diferentes terapias (Guimón 1997). Potencialmente, los ensayos controlados aleatorios o el metaanálisis, (procedimiento en el que se comparan el los tamaños del efecto de las diferentes terapias), son formas útiles de determinar la efectividad diferencial de los diferentes métodos de tratamiento. El metaanálisis permite resumir la información empírica obtenida de un conjunto de investigaciones individuales, y llegar a conclusiones cuantitativas. La condición previa fundamental es la disponibilidad de estudios controlados en los cuales una condición del tratamiento se compara con una condición control, (e.g. grupos de terapia ocupacional, pabellones psiquiátricos ordinarios).

En cuanto a la terapia grupal psiquiátrica, una revisión metaanalítica de McRoberts, Burlingame y Hoag (1998) abarcaba 23 estudios que comparaban directamente la efectividad de los tratamientos de formato individual y grupal..Como la mayoría de las revisiones metaanalíticas anteriores, los resultados de este análisis indicaban que existía escasa diferencia entre la terapia individual y grupal.

La relación desconcertante entre los tamaños del efecto y la *fidelidad* del investigador es un descubrimiento polémico en el metaanálisis (Lambert y Bergin 1994), y recalca la necesidad de una metodología rigurosa para reducir el sesgo en el resultado de la investigación. De esta forma, los métodos de investigación se deberían poder aplicar a diferentes centros de investigación u hospitales de diferentes localidades, (i.e. estudios multi-laterales), para controlar la parcialidad a la fidelidad.

Recientemente, se han tomado iniciativas para mejorar la metodología y el diseño dentro de la investigación de la eficacia de la terapia psicológica. Estas incluyen la introducción de sistemas estandarizados para documentar las características del paciente (e.g. Kallert et

al.2000), así como diversos intentos para estandarizar los instrumentos psiquiátricos que miden los resultados terapéuticos (Piersma y Boes, 1997). Recientemente, se han introducido los sistemas estandarizados de *monitorización a tiempo real* (e.g. el *Sistema de Navegación Sinérgico*, SNS), para la documentación detallada de pacientes y para la evaluación del proceso-resultado a través de teléfonos fijos y móviles (Schiepek et al. 2005). Este tipo de sistema basado en Internet facilita una evaluación rápida, amplia, y frecuente del proceso terapéutico y sus resultados. Sin embargo, en la actualidad, este sistema continua siendo evaluado.

En el pasado, la diversidad y variabilidad de los sistemas de clasificación reflejaban su naturaleza no vinculante. No se había construido una fundamentación teórica. Hasta la fecha, en lo que se refiere a la recogida de información y a la definición de criterios, existe una creciente necesidad de estandarización, al menos en el campo de la investigación. El uso de evaluaciones válidas en los diferentes períodos temporales, (al menos en el pre y postratamiento), y las medidas de seguimiento, se contemplan como indispensables. Como resultado de los desarrollos metodológicos de los criterios de éxito, es posible alcanzar un grado general de acuerdo en varios índices de éxito de tratamiento: el tamaño del efecto, el cambio fiable, y la significación clínica.

Hay tres supuestos fundamentales que constituyen la base de la distinción metodológica entre los niveles de resultado: (i) la decisión preceptiva para la meta del tratamiento de la cura o mejora de los trastornos, (versus otras metateorías de estado psicológicos alterados, e.g., teorías de la etiquetación); (ii) en base a esta decisión, la elección de un modelo genérico de enfermedades como la base común determinada del contenido de las variables resultantes y, (iii) la decisión en torno a una metodología apropiada.

Existe consenso general acerca de que las evaluaciones de resultado se debieran basar en múltiples criterios; debería ser multidimensional, multimetodal, y multimodal. Más aun, se debieran tomar las siguientes medidas en la construcción de un marco de investigación: En un primer paso, el fondo teórico se ha de considerar, para poder así identificar las variables conceptuales, o indicadores a utilizarse en la evaluación del resultado terapéutico. En un segundo paso, se debieran determinar las formas para operativizar y medir las diferentes variables. Esta diferenciación entre aspectos conceptuales y metodológicos es de gran

importancia en la definición de los resultados terapéuticos, ya que cada aspecto es independiente del otro.

2 . Objetivos de la Batería de tests de *Bel- Air*

¿Cuales son las *dimensiones* conceptuales apropiadas para evaluar el éxito de los resultados terapéuticos? Por lo general, el resultado de la terapia se mide en base a los siguientes tres niveles:

- 1) etiología de la disfunción o de las deficiencias
- 2) sintomatología: tanto trastorno específico como general
- 3) roles sociales tales como el “rol del enfermo” o “deterioro del rol normal”.

Hasta ahora, la *etiología* de la disfunción o de las deficiencias se ha pasado, en su mayor parte, por alto en las evaluaciones de los resultados de las intervenciones terapéuticas, (e.g. psicoterapia, farmacoterapia). Según los conocimientos actuales sobre trastornos psiquiátricos, los factores genéticos, bioquímicos, fisiológicos, psicológicos y sociológicos pueden estar operando en esta situación. Todavía, dentro de este contexto, se hace poco uso de la oportunidad de evaluar los parámetros biológicos y neurofisiológicos a través de las técnicas de neuroimagen, tales como PET y el MRI. (cf. Krishnan 1999). Como contraste, algunos de los factores psicológicos involucrados en la permanencia del trastorno (e.g. expresión de las emociones, actitudes cognitivas disfuncionales), se evalúan en los estudios terapéuticos (e.g. NIMH: Estudio para el tratamiento de la depresión: Shea et al. 1992).

La sintomatología se puede abordar de forma directa utilizando inventarios de síntomas tales como: el DSM (Asociación Americana de Psiquiatría, 1994) y el CIE- 10 (Organización Mundial de la Salud, 1992), además de instrumentos de evaluación del observador y de autoinforme validados, que exploren los síntomas psicopatológicos (e.g SCID, SCL-90-R).

El nivel de los roles sociales abraza aspectos tanto de la calidad de vida como aspectos interpersonales o de la interacción social. En los últimos años, se han asentado, en particular, las mediciones para una vida de calidad. Como una dimensión adicional, los *recursos* y las estrategias de afrontamiento adquiridas o “redescubiertas” durante la terapia también podrían tomarse en consideración. Hay un consenso general de que la evaluación de los resultados del tratamiento debiera distinguir al menos estos tres niveles conceptuales, y de que, hasta el momento, se han desatendido las causas de la disfunción.

En las últimas décadas se han realizado diferentes intentos para identificar las medidas que se adecuan a cada uno de los diferentes niveles, para así construir un inventario estandarizado de la evaluación de grupo. La batería de tests desarrollada por la Asociación Americana de Psicoterapia de Grupo es un ejemplo de ello (Dies & Mackenzie 1983). Este inventario estandarizado consiste en una lista de comprobación de síntomas (SCL-90-R), una escala para medir la situación social y los problemas emocionales (Índice de Perfiles Emocionales), una escala para la evaluación global del deterioro psicosocial, (Escala de Evaluación Global e Índice de Mejora General), y un índice de los problemas individuales del paciente durante el proceso terapéutico. Sin embargo, esta batería de tests no recibió una aceptación generalizada, ya que los intereses de los grupos individuales de investigación eran demasiado diversos (Vendar & Kaul, 1994).

En base a estas observaciones y a la situación actual de debate en la investigación psicoterapéutica, se han formulado las siguientes recomendaciones para una evaluación sólida de los resultados terapéuticos (Straub et al., 1996):

- La metodología aplicada debe permitir comparar los resultados de diferentes formas de tratamiento, (psicoterapéutico y farmacoterapéutico).
- -La investigación evaluativa debe evitar infringir sobre el tratamiento clínico.
- Las características del paciente deben documentarse de la forma más precisa posible, haciendo uso de métodos estandarizados (e.g. sistemas establecidos de documentación base y clasificación diagnóstica).
- Se deben utilizar Instrumentos consolidados, (i.e. validados), que intervengan a varios niveles del cambio potencial (causas y deficiencias, síntomas y roles sociales).
- Los efectos negativos también se deben registrar con detalle (e.g. características de abandono).

3. . La Bateria de Tests de *Bel Air* para la Evaluación de Trastornos Psiquiátricos

Los departamentos psiquiátricos de hoy en día implican una variedad de escenarios de tratamiento para las intervenciones psicosociales, (e.g. unidades de hospitalización, unidades para pacientes externos, clínicas de día, departamentos para la atención primaria de pacientes con síntomas crónicos). Para la evaluación de los resultados terapéuticos, se necesitan un conjunto común de herramientas de medida, que se puedan utilizar en cada uno de estos

escenarios y con varios grupos de pacientes psiquiátricos. Las pruebas de valoración rápida constituyen unas herramientas válidas, poderosas y económicas para satisfacer esta demanda. Para poder evaluar la efectividad de las diferentes terapias desarrolladas en el Departamento de Psiquiatría de la Universidad de Ginebra, se evaluaron los resultados de los diferentes métodos de tratamiento a varios niveles, (ver sección 2): en una evaluación global de los resultados terapéuticos por pacientes y terapeutas, y en una evaluación más diferenciada de los cambios (pre-post tratamiento) de las respuestas de los pacientes a cuestionarios que exploraban aspectos y síntomas de los roles sociales. Una visión global de los instrumentos se da en la Tabla 1.

Tabla 1: Niveles conceptuales y aplicación metodológica de un programa de evaluación para diferentes trastornos psiquiátricos. (Bateria de tests de la Universidad de Ginebra).

<u>Nivel de Evaluación</u>	<u>Variables</u>	<u>Instrumentos</u>
Síntomas	-Quejas/psicopatología (auto-informe) -Diagnóstico(índices clínicos)	-Inventario Breve de Síntomas (IBS, Derogatis & Melisaratos 1983; IBS, Derogatis 1993) -DSM-IV (Asociación Americana de Psiquiatría, 1994).
Roles sociales	-Relaciones personales (Auto-informe) -Funcionamiento social (índices clínicos) -Funcionamiento social (Auto-informe).	-Cuestionario de Evaluación de las Relaciones (ERA, Fredenrich & Zinetti, 2000) -Evaluación Global del Funcionamiento (GAF, DSM-IV-Eje V, Asociación Americana de Psiquiatría, 1994) -Cuestionario de las Funciones Sociales (QFS, Zanello et al. 2006)

A continuación, se presentarán los elementos centrales de las baterías de tests- IBS, GAF, ERA y QFS, y se expondrán las razones por las que se eligieron estos instrumentos en particular.

Maercker & Menning 18

El Inventario Breve de Síntomas, (IBS, Derogatis & Melisaratos, 1983; Derogatis, 1993) es una medida de autoinforme breve apropiada para la evaluación de la psicopatología y las quejas psicológicas. Es una versión reducida de la SCL-90-R, diseñada por los mismos autores, y que incluye sólo 53 de los 90 ítems originales. De esta forma, el tiempo de aplicación se recortó de 15- 20 minutos a 5-10 minutos. Los ítems se resumen para formar escalas de 9 síntomas: (somatización, obsesivo-compulsivo, depresión, ansiedad, ira-hostilidad, ansiedad fóbica, ideación paranoide, sensibilidad interpersonal y psicoticismo), y tres índices globales: (índice de gravedad global, índice de angustia de síntomas positivos, y total de síntomas positivos).

Hasta ahora, el IBS se ha utilizado básicamente para estudios con pacientes afectados por trastornos psicológicos, (para una visión general ver Piersma et al. 1994), y en la evaluación de los resultados de tratamientos psico/farmacoterapéuticos, (para una visión general ver Franke, 1997). El IBS se ha traducido ahora a varios idiomas, (e.g. castellano, alemán, hebreo, italiano y francés).

La Evaluación del Funcionamiento Global (GAF) se seleccionó para evaluar los roles sociales y el funcionamiento social de forma general, utilizando una única medida. El GAF equivale al eje V del DSM-IV (Asociación Americana de Psiquiatría, 1994). Fue elaborado y operativizado por Endicott et al. (1976), como una escala índice de la enfermedad global. La escala GAF se debe valorar únicamente con respecto al funcionamiento psicológico, social y ocupacional. Las instrucciones excluyen el deterioro del funcionamiento social debido a limitaciones físicas, (o ambientales). La escala GAF considera estas funciones dentro de un continuum hipotético de salud-enfermedad mental. El GAF puede valorarse en varios periodos temporales, (e.g. actual o mayor grado de funcionamiento durante al menos un mes en el pasado año).

El Cuestionario para la Evaluación de las Relaciones (French: ERA, Fredenrich & Zinetti, 2000) es una nueva medida de autoinforme que se pretende sea igualmente apropiada para los diferentes trastornos psiquiátricos y psicoterapéuticos,(e.g. esquizofrenia, depresión y trastornos ansiosos), así como más corta y económica que las herramientas existentes. El ERA consiste de 16 ítems, capturando 4 dimensiones (=factores en el análisis factorial): 1.) Abertura hacia los otros, 2.) Distancia relacional, 3.) Percepción relacional/personal y 4.) Placer en/a través de las relaciones (Zinetti-Bertschy et al. 2004). La ERA muestra una

elevada validez convergente, una relación sistémica con aspectos clínicos, (con el IBS $r=0.6$, y QFS $r=0.5$), una elevada sensibilidad al cambio, y una buena utilidad por gracias a su brevedad y su rápida y práctica ejecución.

El Cuestionario de las Funciones Sociales, ((French QFS, Zanello et al. 2006) era una herramienta breve y económica utilizada para evaluar las funciones sociales en diferentes trastornos psiquiátricos. La Escala de Ajuste Social, (SAS: Weissman & Bothwell 1976) no fue seleccionada porque se consideró demasiado extensa y redundante en un estudio piloto. El QFS engloba un total de ocho campos, con dos preguntas pertinentes a cada campo: (1. frecuencia de la actividad 2. satisfacción de la actividad): 1. Actividades 2. Actividades de cada día 3. Ocio 4. Familia y relaciones de pareja 5. Relaciones extrafamiliares 6. Gestión administrativa y financiera 7. Salud general, y 8. Vida social e información. Existe un estudio de validación que halló propiedades psicométricas satisfactorias para el QFS. Esta escala posee cualidades equivalentes o mejores en comparación con otras herramientas validadas, [como el SAS, La Escala de Adaptación Social y Auto Evaluación, o el SAS-SR (validez convergente ($r=0.65$)), y el Autoinforme del Ajuste Social]. Parece ser un instrumento valioso para la supervisión de las funciones sociales en los pacientes psiquiátricos porque: las normas clínicas están disponibles, revela una elevada sensibilidad al cambio, (estudiada en pacientes externos) y es más corta, más rápida, y de una mayor fiabilidad en contraste con herramientas anteriores.

Todas las escalas resumidas en el presente documento constituyen el fundamento de los instrumentos de medida del resultado para varias poblaciones psiquiátricas. Durante la última década, los abordajes a trastornos específicos han ganado una mayor atención. En lo que queda del capítulo, nos gustaría destacar el campo del Trastorno por Estrés Postraumático (TEPT), y estados relacionados, que son el núcleo de nuestra experiencia clínica e investigadora (e.g. (e.g., Maercker, Zöllner, Menning, Rabe, Karl, 2006; Knaevelsrud & Maercker, 2007). No sólo son los pacientes TEPT los que necesitan de una evaluación apropiada tras el evento traumático o por sintomatología TEPT, sino también una variedad de otros pacientes que padecen de otros trastornos (trastornos afectivos, trastornos de ansiedad, y trastornos de personalidad límite).

4. Evaluación de Trastornos Agudos y por Estrés Postraumático en la intervención en crisis.

La sintomatología del estrés postraumático se aborda en el primer mes como un trastorno por estrés agudo (TEA), cumpliendo con los criterios A del DSM-IV, (el evento traumático), adhiriéndose a la mayoría de los síntomas de TEPT, y, además, al criterio de disociación. Sin embargo, la característica distintiva del TEPT es la persistencia de síntomas como la intrusión, la hipervigilancia, la evitación, y el entumecimiento durante un tiempo superior a un mes.

Pero, no es tan sólo relevante la distinción entre TEPT y TEA en la práctica clínica, sino que también es posible que lo sea en vertientes del desarrollo postraumático como el Trastorno de Adaptación (TA), Trastorno del Duelo Complicado (CGD Forstmeier and Maercker, 2007, Maercker, Einsle and Koellner, 2007), TEPT complejo, e incluso, en Trastornos de la Personalidad Límite, (TPL). Para estas evaluaciones, se presentan dos instrumentos breves: PTSS-10 y la escala para el TEPT de Breslau.

pgs. 231 y 232:

La Escala de Estrés Postraumático (*The Posttraumatic Symptoms Scale PTSS-10*) es un instrumento de 10 ítems concebido de acuerdo con el DSM. Fue desarrollada inicialmente en Noruega, después de una catástrofe en una plataforma petrolífera (Weisaeth, 1989). Está disponible en 8 idiomas, y es muy apreciada porque se puede aplicar en menos de tres minutos. Resulta especialmente útil para los supervivientes de catástrofes graves y de torturas. Holen (1990) utilizó una versión unifactorial para la población noruega.

Sin embargo, su sensibilidad es bastante baja (75%), para un umbral de 35, debido a una excesiva inclusión de síntomas depresivos. En contraste, su especificidad es muy alta (98%), así como su valor predictivo positivo y negativo (91 y 93%). La consistencia interna (Alpha Conbachs) y su fiabilidad test-retest son bastante buenas (0.85 y 0.89, Holen, 1990). Actualmente, el PST-10 es ampliamente utilizado en las Unidades de Urgencia, y en crisis médicas y psiquiátricas para explorar el estrés postraumático.

La Escala de despistaje de Estrés Postraumático de Breslau, (*The Breslau PTSD Screening Scale* (Breslau et al., 1999)) es un instrumento nuevo aún más reducido. Consiste en sólo 7 ítems elegidos entre los criterios más sensibles y específicos del DSM IV. Su sensibilidad y especificidad son bastante buenas (80% y 97% respectivamente). Hoy día es una herramienta

ampliamente utilizada en los EE.UU. No incluye síntomas intrusivos debido a su baja especificidad, por lo que para su uso clínico hemos añadido dos ítems intrusivos que cubran todos los síntomas del TEPT. Esta versión, es tanto en la clínica como en la investigación, un método eficaz para el despistaje del TEPT.

5. Perspectivas

Como se ha mostrado en la revisión anterior, se ha realizado un gran progreso en la evaluación metodológica. En particular, la elaboración de instrumentos breves y económicos ha marcado un hito en el desarrollo de los últimos años; la aplicación es más fácil tanto para el paciente como para el evaluador, y existe mejor ajuste al día a día de la práctica clínica.

Los métodos de evaluación continuarán desarrollándose en los años siguientes. Se puede esperar que incluso en terapias no-biológicas, (no farmacológicas), se hagan uso de métodos biopsiquiátricos, para evaluar las causas de la disfunción. Tales métodos biopsiquiátricos incluyen parámetros para la expresión del gen, neurotransmisores, y psicofisiología, (e.g. riego cerebral; Crespo-Facorro et al. 1999). Sin embargo, dentro del campo de la psiquiatría, las causas psicológicas de la disfunción y la deficiencia continuarán desempeñando un papel fundamental, (los parámetros de la expresión de la emoción, las tergiversaciones cognitivas de síntomas básicos, las soluciones a problemas interpersonales disfuncionales).

Otro nivel conceptual de evaluación posible, (de la Tabla 1), es el nivel social de valores culturales compartidos (Guimon et al. 1999; Maercker & Müller 2000). Este aspecto es de particular relevancia con respecto a la cuestión de si los métodos terapéuticos de base empírica son utilizables en el tratamiento de trastornos psicológicos, en países con recursos financieros mas limitados.