

Tools for Assessment

Prepublication long form. A shorter form of this chapter is in the

Gazzillo, F., Gordon, R.M. & Waldron, S. (2017). Tools, Chapter 6, in
Psychodynamic diagnostic manual, Version 2 (PDM-2). New York:
Guilford Press.

Section's Editors

Francesco Gazzillo, PhD, "Sapienza" University of Rome, Italy

Robert M. Gordon, PhD, ABPP, Allentown, Pennsylvania, USA

Sherwood Waldron, MD, Psychoanalytic Research Consortium, New York

Consultants¹

Alan Abbass, from Halifax, Nova Scotia, USA
John Auerbach, from Gainesville, Florida, USA
Jacques Barber, from Garden City, NY, USA
Mark Blais, from Boston, Massachusetts, USA
Anthony Bram, from Lexington, Massachusetts USA
Robert Bornstein, from Garden City, New York, USA
Manfred Cierpka, from Heidelberg, Germany
John Clarkin, from White Plains, New York, USA
Franco Del Corno, from Milan, Italy
Kathryn DeWitt, from Stanford, California, USA
Diana Diamond, from New York City, USA
Federica Genova, from Rome, Italy
Carol George, from Oakland, CA
Mark Hilsenroth, Adelphi University, Long Island, NY, USA
Per Høglend, from Oslo, Norway
Steven Huprich, from Wichita, Kansas, USA
Lawrence Josephs, from Garden City, New York, USA
Horst Kaechele, from Berlin, Germany
Guenther Klug, from Munich, Germany
Michael Lambert, from Provo, Utah, USA
Patrick Luyten, from Brussels, Belgium
Marianne Leuzinger-Bohleber, from Frankfurt, Germany
Seymour Moscovitz, from New York City, New York, USA
John Norcross, from Clarks Summit, Pennsylvania, USA

¹ The consultants' contributions have been invaluable. The editors wish to give particular acknowledgement to the extensive contributions to the whole manuscript by John Auerbach.

Chris Perry, from Montreal, Canada

Piero Porcelli, Castellana Grotte (Bari), Italy

John Porcerelli, from Bloomfield Hills, Michigan, USA

Jeremy Safran, from New York City, NY, USA

George Silberschatz, from San Francisco, California, USA

Annamaria Speranza, from Rome, Italy

Michelle Stein, from Boston, Massachusetts, USA

William Stiles, from Miami, Ohio, USA

Karl Stukenberg, from Cincinnati, Ohio, USA

Annalisa Tanzilli, from Rome, Italy

Introduction

The aim of this section is to survey a broad range of empirically validated assessment tools that may enhance your use of the PDM-2 for diagnostic formulations, clinical reports, treatment planning, or research. We also describe several ways that these various measures have already led to findings relevant to psychopathology and psychotherapy.

First, we will introduce two specific PDM-2 derived assessment tools: the Psychodiagnostic Chart-2 (PDC-2) and the Psychodynamic Diagnostic Prototypes-2 (PDP-2). The Psychodiagnostic Chart-2 is aimed at guiding the clinician through the entire PDM-2 diagnostic formulation: personality organization, personality patterns and disorders, mental functioning, subjective experience of symptoms, manifest symptoms and complaints, and cultural, contextual and other relevant considerations. The original Psychodiagnostic Chart (PDC; Gordon & Bornstein, 2012) was derived from the first edition of the PDM and was used and validated both in public and private settings in US and Europe.

The Psychodynamic Diagnostic Prototypes (PDP; Gazzillo, Lingardi, Del Corno, 2012) provides prototypic examples of the various personality styles or disorders listed in the P Axis (adult) of the PDM-1, together with prototypical descriptions of the different levels of personality organization and the main features of the anaclitic and introjective psychologies and personality disorders. The PDP-2 (Gazzillo, Genova, & Lingardi, 2016) allows a more precise assessment of the patient according to the P Axis dimensions.

After introducing these two specific PDM-2 tools, we describe additional instruments that can enhance the assessment of patients according to the PDM-2 (personality organization, personality patterns or disorders, and mental functions). We will tend to avoid tests of specific symptoms or syndromes out of the context of personality dynamics.

Some of these clinician-rated instruments directly influenced the construction of the first and second editions of the PDM. These include the Shedler-Westen Assessment Procedure (SWAP; Westen & Shedler, 1999a, 1999 b; Westen, Shedler, Durrett, Glass, Martens, 2003; Westen, Shedler, Bradley & DeFife, 2012), the Structural Interview for

Personality Organization (STIPO; Stern et al., 2009), the Karolinska Psychodynamic Profile (KAPP; Weinryb, Rossel, & Asberg, 1991), the Defense Mechanisms Rating Scale (DMRS; Perry, 1990), the Social Cognition and Object Relations Scale (SCORS; Westen, 1991a, 1991b; Stein et al., in press), and the Object Relations Inventory (ORI; Blatt, & Auerbach, 2001).

The other instruments reviewed here, such as self-reports (e.g., Minnesota Multiphasic Personality Inventory-2 [MMPI-2]; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), and performance-based assessment tools (e.g., the Rorschach Inkblot Method [Rorschach, 1921/1942] or the Thematic Apperception Test [Murray, 1943]) are widely used in both assessment practice (Bram & Peebles 2014) and psychodynamic research (Bornstein, 2010) and can be a useful aid in a PDM-2 formulation for describing and understanding the psychological capacities and personality features of patients and their overall level of functioning and well-being.

Lastly, we describe some useful instruments, not necessarily directly related to PDM-2 case formulation, that are valuable nevertheless in the empirical assessment of the processes of psychodynamic psychotherapy and psychoanalyses. These instruments assess patients' psychological and mental capacities as they appear during therapy sessions, and consequently may contribute to the assessment of patients' functioning through direct observation in this interpersonal context.

General Considerations

Personality assessment can be considered from different perspectives. One of these is the point of view from which the data are derived: (a) the clinician's point of view (clinician report) (like SWAP and STIPO); (b) the patient's point of view (self-report) (like MMPI-2); (c) significant others' reports of the patient; (d) an independent judge or observer of the patient; and (e) the patients' performance, often on unstructured or minimally structured tasks that were formerly described as projective.

Another perspective on assessment is the content or focus of the assessment: (a) the

patient's internal personality structure, (b) the patient's external observable behavior in daily interactions, the patient's ideas about his or her personality features, level of perceived wellness, and capacity of functioning, and (c) the patient's observable behavior within the consulting office. We would add another: measures of the patient contributions to the psychotherapeutic process.

Research has shown considerable divergence among these points of view and methodologies (Bornstein, 2002, 2009; Hunsley & Mash, 2014). Similarities and differences between the results of assessments developed from the different points of view and methods (clinician-report, self-report, other-report, or performance based) may be quite informative from a clinical perspective (Bram & Yalof, 2014; Bram & Peebles, 2014), with respect both to diagnostic formulation and to conceptualization of personality dynamics and interpersonal functioning (Bornstein, 2010). For example, a large divergence between the patient's self-report and results of a performance-based test may indicate the patient's difficulty in mentalizing, the influence of defensive processes blocking self-awareness, or dissimulation on the part of the patient. It is well known that unconscious motivations, more easily assessable with performance-based instruments, are more predictive of long-term life choices than are conscious motivations accessible through self-reports (Westen, 1999). Emotional reactions not recognized consciously, and hence not easily accessible through self-report measures, influence both behavior and physical health, and can be identified with performance-based methods, physiological data or well-validated clinician-report tools (Shedler, Mayman, & Manis, 1993).

Clinicians observe patients and interact with them, ask them to describe themselves, their lives, their relationships, and what brings them to assessment or treatment and are sensitive to the emotions that patients stir up in them. These are unique data that need a clinician's perspective to interpret, even though such interpretation tends to be highly variable (e.g. Seitz 1966). The variability is decreased when these descriptions are systematized with sound empirical tools, as described below. Validated clinician-report instruments may help a clinician to perform this task by supporting, amplifying, or modifying clinical impressions and self-report or performance-based data in a reciprocally

enriching process (Meehl, 1996; Westen & Weinberger, 2004).

Both assessment validity and clinical utility are improved by using and integrating a battery of assessment instruments from different points of view and methodologies (e.g., self-report, expert ratings, performance-based procedures). The chapters in Hopwood and Bornstein (2014) provide guidelines for integrating multi-method assessment data in different areas (e.g., psychopathology, resilience, risk management), and Bornstein (2010) provides an overarching conceptual framework for integrating multi-method assessment data from a psychodynamic perspective.

The selection of instruments in any assessment should be based on the goals of that assessment. Clinicians may be interested in using a variety of tools to tap different aspects of personality for a more complete diagnostic formulation and treatment plan. Or the goal may be to evaluate the impact of treatment, in which case, the clinician may prefer measures that best reflect personality changes. Regardless of the goal of the assessment, instruments with contrasting methodologies should be employed, if feasible, to illuminate convergences and divergences between different perspectives and methods (Bornstein, 2009; Cogswell, 2008; De Fife et al., 2010)

Many of the tools we will describe are also time consuming, and this feature will often discourage their use in routine clinical practice. Nevertheless, their richness of information and their methodological and theoretical soundness can prove valuable to clinicians at all levels of experience, and therefore, especially when complex clinical questions are being addressed, implementation of more time-consuming, labor-intensive assessments is warranted.

Bram and Yaloff (2014) discuss three particularly useful clinical issues that benefit from testing: (a) assessment of suitability for treatment, (b) evaluation of treatments that are not proceeding well, and (c) modification of treatment to take account of aspects of the personality (including limitations) that influence, positively or negatively, the likelihood of treatment success.

Most of the measures we have chosen are aimed at the assessment of personality organization, fundamental mental functions, and global adjustment and well-being and not

simply at symptom patterns or symptom severity. We have made these choices because symptoms such as alteration in sleep and eating, depressed mood, inexplicable pain, sexual difficulties, phobias, or panic attacks are easily observable by clinicians or explicitly reported by patients or both. In contrast, personality features and implicit or unconscious mental processes are not. Additionally, from a psychodynamic perspective, the assessment of personality organization and features is relevant for understanding the meaning, function, etiology, and prognosis of the different syndromes and symptoms (Westen, Gabbard, & Blagov, 2009; Westen & Harnden-Fisher, 2001; Thompson-Brenner & Westen, 2005a; Thompson-Brenner & Westen, 2005b; Thompson-Brenner, Eddy, Satir, Boisseau, & Westen, 2008; Gazzillo, Lingiardi, Peloso, et al., 2013).

Finally, virtually all the instruments we describe provide both dimensional and categorical assessment, in line with current clinical and research thinking.

We will provide a description of each of the instruments we have chosen and their main strengths and limitations. Because a more complete review of these assessment tools is beyond the purpose of this section, references are provided at the end of each tool description to encourage the reader to learn more about the instruments before using them.

In the first part of this section (below) is on clinician reports. We begin by describing measures that assist in PDM-2 diagnosis of personality organization and mental functioning, and are easiest for clinicians to use. Then we discuss additional tools that require more training for the clinician or clinician-researcher but that are important for clinicians to know about because of the richness of empirical findings relevant to clinical work. We have arranged the second part, on self-reports, so that the order of the tools described corresponds with the PDM-2 tasks of first determining level of personality structure, then personality patterns, and finally mental functioning.

References

Blatt, S. J., & Auerbach, J. S. (2001). Mental representation, severe psychopathology, and the therapeutic process: Affect and self-reflexivity in borderline and schizophrenic

- patients. *Journal of the American Psychoanalytic Association*, 49, 113–159.
- Bornstein, R. F. (2002). A process dissociation approach to objective-projective test score interrelationships. *Journal of Personality Assessment*, 78, 47-68.
- Bornstein, R. F. (2009). Heisenberg, Kandinsky, and the heteromethod convergence problem: Lessons from within and beyond psychology. *Journal of Personality Assessment*, 91, 1-8.
- Bornstein, R. F. (2010). Psychoanalytic theory as a unifying framework for 21st century personality assessment. *Psychoanalytic Psychology*, 27, 133-152.
- Bram, A.D. and Peebles, M.J. (2014). *Psychological Testing that Matters: Creating a Road Map for Effective Treatment*. American Psychological Association, Washington, DC.
- Bram, A.D & Yalof, J. (2014) *Quantifying Complexity: Personality Assessment and its Relationship with Psychoanalysis*. *Personality Assessment*, In Press.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2: Manual for administration and scoring*. University of Minnesota Press, Minneapolis, MN.
- Cogswell, A. (2008). Explicitly rejecting an implicit dichotomy: Integrating two approaches to assessing dependency. *Journal of Personality Assessment*, 90, 26-35.
- DeFife, Jared A; Drill, Rebecca; Nakash, Ora; Westen, Drew. (2010). Agreement between clinician and patient ratings of adaptive functioning and developmental history. *The American Journal of Psychiatry*, 167: 1472-1478.
- Gazzillo, F., Lingiard, V., & Del Corno, F. (2012). Towards the validation of three assessment instruments derived from the PDM Axis P: The Psychodynamic Diagnostic Prototypes, the Core Preoccupations Questionnaire and the Pathogenic Beliefs Questionnaire. *Bollettino di Psicologia Applicata*, 265, 1–16.
- Gazzillo, F., Lingiard, V., Peloso, A., Giordani, S., Vesco, S., Zanna V., Filippucci, L. & Vicari, S. (2013): Personality subtypes in adolescents with eating disorders. *Comprehensive Psychiatry*, 54, 6, 702-712.
- Gazzillo, F., Genova, F. & Lingiard, V. (2016). *Psychodynamic Diagnostic Prototypes -2 (PDP-2)*. Unpublished Manuscript. Department of Dynamic and Clinical Psychology,

“Sapienza” University of Rome.

Gordon, R. M., & Bornstein, R. F. (2012). The Psychodiagnostic Chart (PDC): A practical tool to integrate and operationalize the PDM with the ICD or DSM. Retrieved from [http:// www.mmpi-info.com/pdm-blog](http://www.mmpi-info.com/pdm-blog)

Hopwood, C.J. & Bornstein, R.F. (2014) (Eds) *Multimethod Clinical Assessment*, Guilford Press, New York.

Hunsley, J., & Mash, E. J. (2008). *A guide to assessments that work*. Oxford University Press, USA.

Hunsley, J., & Mash, E. J. (2014) "Evidence based assessment", *Oxford handbook of clinical psychology* (rev. ed.), D. H. Barlow (Ed.), New York, Oxford University Press, 76-97

Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical–statistical controversy. *Psychology, Public Policy, and Law*, 2, 293-323.

Murray, M.A. (1943) *Thematic apperception test*. Cambridge, MA, US: Harvard University Press

Perry, J. C. (1990). *Defense Mechanism Rating Scale (DMRS)* (5th ed.). Boston, MA: Cambridge University.

Rorschach, H. (1942). *Psychodiagnostics: A diagnostic test based on perception* (P. Lemkau & B. Kronenberg, trans.). Berne, Switzerland: Huber. (Original work published 1921)

Seitz, P. (1966). The Consensus Problem in Psychoanalytic Research. In *Methods of Research in Psychotherapy*. Gottschalk, L. & Auerbach, A. H. (Eds.). New York: Appleton-Century-Crofts.

Shedler J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, 48, 1117–1131.

Stern B.L, Caligor E, Clarkin J.F, Critchfield K.L, MacCornack V, Lenzenweger M.F, Kernberg O.F (2010):The Structured Interview of Personality Organization (STIPO): Preliminary Psychometrics in a Clinical Sample. *Journal of Psychological Assessment*,

91, Vol. 1, 35-44.

- Thompson-Brenner H., Westen D. A (2005a): A naturalistic study of psychotherapy for bulimia nervosa. *Journal of Nervous and Mental Disease*; 193:573–595
- Thompson-Brenner H, Westen D. (2005b): Personality subtypes in eating disorders: Validation of a classification in a naturalistic sample. *British Journal of Psychiatry*;186:516–524
- Thompson-Brenner H., Eddy K.T., Satir D.A, Boisseau C.L., Westen D. Personality subtypes in adolescents with eating disorders: Validation of a classification approach. *Journal of Child Psychology and Psychiatry*.
- Weinryb R.M., Rössel R.J. (1991): Karolinska Psychodynamic Profile - KAPP. *Acta Psychiatrica Scandinavica*, 85, 153–162
- Westen, D. (1991a). Clinical assessment of object relations using the TAT. *Journal of Personality Assessment*, 56, 56-74.
- Westen, D. (1991b). Social cognition and object relations. *Psychological Bulletin*, 109, 429-455.
- Westen, D. (1999). The Scientific status of unconscious processes: is Freud really dead?. *Journal of the American Psychoanalytic Association*, 47, 1061-1106.
- Westen, D., Gabbard, G. O., & Blagov, P. (2006). Back to the future: Personality structure as a context for psychopathology. In R. F. Krueger & J. L. Tackett (Eds.), *Personality and psychopathology* (pp. 335–384). New York, NY: Guilford Press.
- Westen, D., Harnden-Fischer, J. (2001): Personality profiles in eating disorders. In Leary, M. R., Tangney, J.O.: *Hanbook of self and identity*. Guilford Press, New York, 643-664.
- Westen, D., & Shedler, J. (1999a). Revising and assessing axis II, Part I: Developing a clinically and empirically valid assessment method. *The American Journal of Psychiatry*, 156, 258–272.
- Westen, D., & Shedler, J. (1999b). Revising and assessing axis II, Part II: Toward an empirically based and clinically useful classification of personality disorders. *The American Journal of Psychiatry*, 156, 273– 285.

- Westen, D., Shedler, J., Durrett, C., Glass, S., & Martens, A. (2003). Personality diagnosis in adolescence: DSM–IV Axis II diagnoses and an empirically derived alternative. *American Journal of Psychiatry*, 160, 952–966.
- Westen, D. & Weinberger, J. (2004). When clinical description becomes statistical prediction”. In *American Psychologist*, 59, 595-613.
- Westen, D., Shedler, J., Bradley, B., & DeFife, J. A. (2012). An empirically derived taxonomy for personality diagnosis: Bridging science and practice in conceptualizing personality. *The American Journal of Psychiatry*, 169, 273–284.

Part One: Systematic Evaluation by the Clinician

A. PDM-2 Instruments

We start with our specifically developed tools for the PDM/PDM-2 and then review assessments that can be helpful to your PDM-2 diagnostic formulation.

In 2011, Robert M. Gordon and Robert F. Bornstein agreed that the PDM needed a short, user-friendly tool that would: (a) guide the practitioner through all sections of the PDM taxonomy, (b) be idiographic, flexible and useful for clinicians of any theoretical orientation, and (c) integrate the PDM with the symptom classifications of the Diagnostic and Statistical Manual (DSM) or International Classification of Diseases (ICD). We developed the PDC with this in mind.

The PDC-2 is an updated version of the PDC that is compatible with the PDM-2 and that incorporates the changes that were introduced during the PDM revision process. The modifications are not expected to affect reliability and validity of the PDC-2. The PDC-2 is for use with adult clients, but the basic format of the PDC-2 can be used for the Elderly, Adolescent, Child and Infancy sections or further revisions of the PDM.

The goal of the PDC-2 is to offer a person-based nosology that may be used for teaching the PDM-2, supervision, diagnoses, treatment formulations, progress reports, and outcome assessment, as well as for empirical research on personality, psychopathology, and treatment. Our overarching aim is to make psychodiagnostic formulation more useful to the practitioner by combining the symptom-focused ICD and DSM with the full range and depth of human mental functioning addressed by the PDM (see, e.g., Bornstein, 2011; Lingiardi et al., 2015).

The PDC-2 requires some familiarity with the PDM-2. The clinician must perform (or have access to) diagnostic interview data and psychological assessment data to derive optimal PDC-2

ratings. We recognize that this is not always feasible, and in many instances the clinician will code an initial impression, then reassess as additional information accrues. If the PDC-2 is used for progress notes, there will be opportunities to reassess and revise the person's diagnosis as well. The validity of this chart can be enhanced with the integration of relevant psychological tests as suggested in this chapter.

Qianna Snooks helped to format the PDC-2 v.8, as well as a digital form for computers that has a reset button for repeated uses, has a drop-down menu for dominant personality patterns, automatically totals the individual M axis scores, and has a large dialog box for further contextual information. The digital form also allows the ratings to be exported to Excel for data collection and statistical analysis for research purposes.

Initially, to assess the utility of the PDC, Gordon, Bornstein and Stoffey recruited expert assessment practitioners from various psychology listservs and websites to complete an online survey after using the PDC with at least one client. The survey asked how the practitioners valued various PDM-PDC sections, as well as the DSM and ICD symptom-based nosologies. We analyzed the data when we had 50 completed surveys. Half the respondents identified themselves as Psychodynamic (50%); the rest were Eclectic (22%), Cognitive-Behavioral (12%), Humanistic/Existential (10%), Systems (4%), and Other (2%). Eighty percent of this sample had doctoral degrees.

Sixty-eight percent of the practitioners rated the PDM Personality Structure (subsequently labeled "Personality Organization") as "helpful-very helpful," 58% rated PDM Mental Functioning as "helpful-very helpful," and 44% rated PDM Dominant Personality Patterns or Disorders as "helpful-very helpful." In contrast, only 18% of practitioners rated DSM Global Assessment of Function (GAF; American Psychiatric Association, 2000) scores as "helpful-very helpful," and 14% rated ICD or DSM symptoms as "helpful-very helpful." These preliminary results lend strong support for the PDM's taxonomy and the usefulness of the PDC (Bornstein & Gordon, 2012; Gordon & Stoffey, 2014).

A follow up study by Gordon, Blake, Bornstein, Gazzillo, Etzi, Lingiardi, McWilliams, Rothery, and Tasso (2015) asked the same survey questions to a sample who were not assessment experts but more “typical” mental health practitioners from a wide range of educational backgrounds and theoretical orientations ($N= 438$). They were asked to diagnose a recently seen patient and then rate how helpful various personality diagnostic taxa were in understanding their patient. In this sample, 46% held doctoral degrees, and theoretical orientations were 26% Psychodynamic, 33% CBT, and 41% Other (e.g., Family Systems, Humanistic-Existential, Eclectic and other).

The results of our survey indicated that the percent rated as “helpful – very helpful” in understanding their patients for each of five diagnostic taxa were Level of Personality Organization 75%, Personality Disorders 62%, Mental Functioning 67%, and Cultural/Contextual Dimension 41%. Only 30.5% rated Symptoms as “helpful-very helpful” in understanding their patient. All differences were statistically significant.

These studies support the clinical utility and structure of the PDC/PDC-2 with experts who are mainly psychodynamic, as well as with “typical” practitioners of various theoretical orientations.

Gordon and Stoffey (2014) found that the PDC has excellent internal consistency and good test-retest stability over two weeks (ranging from $r = .69$ to $.92$). They also reported excellent construct validity for the PDC when compared to the scales of the MMPI-2, the Operationalized Psychodynamic Diagnosis (OPD) Axis IV Psychic Structure/Mental Functioning scales, and the scales of the Karolinska Psychodynamic Profile (KAPP).

Overall, there is sufficient evidence at this time of the reliability, validity and practical utility to support use of the PDC in clinical and research settings. These results are generalizable to the PDC-2. The PDC has been shown to be a valuable instrument for personality research (Gordon, Stoffey, & Perkins, 2013; Huprich, Lingiardi, McWilliams, Bornstein, Gazzillo, & Gordon,

2015).

(For a full presentation of these results, copies of the PDC-2 and instructional manual, search online “Psychodiagnostic Chart-2”).

The PDC-2 starts with Section I, the Level of Personality Organization. The Personality Organization scale is a 1-10 rating in order to incorporate ranges within the four categories: Psychotic, Borderline, Neurotic, and Healthy levels.

The practitioner rates four mental functions on a scale from 1 (severely impaired) to 10 (high functioning): (a) *Identity*, the ability to view the self in a complex, stable, and accurate ways; (b) *Object relations*, the ability to maintain intimate, stable, and satisfying relationships; (c) *Level of Defenses*; and (d) *Reality testing*, the ability to appreciate conventional notions of what is realistic. Then the practitioner rates the client’s overall Personality Organization (psychotic, borderline, neurotic, or healthy) on the basis of his/her ratings of the four component scales and overall clinical judgment.

Section II asks the practitioner to determine the client’s Personality Patterns or Disorders, (P-Axis) by checking as many relevant personality patterns as apply.; then the practitioner notes the one or two most dominant patterns. For research purposes, each personality pattern can be given a rating from 1 (severe impairment) to 5 (high functioning).

Section III asks the practitioner to determine the client’s Mental Functioning (M-Axis), This is a detailed consideration of the patient’s various strengths and limitations along 12 dimensions that are subsumed in four main domains: cognitive and affective processes, identity and relationships, defense and coping, and self awareness and self direction. The 12 mental functions are summed to derive an overall Level of Severity score, as follows: Healthy/Optimal Mental Functioning 54-60; Appropriate Mental Functioning with Some Areas of Difficulty 47-53; Mild Impairments in Mental Functioning 40-46; Moderate Impairments in Mental Functioning 33-39;

Major Impairments in Mental Functioning 26-32; Significant Defects in Basic Mental Functions 19-25; Major/Severe Defects in Basic Mental Functions 12-18. The digital form of the PDC-2 automatically derives these scores.

Section IV asks the practitioner to describe the main symptom patterns from the S-Axis (e.g., those that are related to psychotic disorders, mood disorders, anxiety disorders, event and stress disorders, specific symptom disorders, addiction and medically related disorders, etc.) If required, the practitioner may use DSM or ICD symptoms and codes here. The dominant symptoms are rated on a 5-point scale from 1 (severe) to 5 (mild).

In Section V, the practitioner may choose to add relevant information regarding Cultural, Contextual and Other Relevant Considerations. The digital form of the PDC-2 has a dialog box allowing for more narrative space.

Finally, the practitioner integrates all the sections into a cohesive clinical picture that provides a nuanced understanding of the whole person.

Strengths: The PDC has good reliability and validity. It guides the practitioner through a PDM-2 formulation using all the axes. It is only three pages long and user friendly. It provides a broad summary of the PDM-2 Adult Axis and can be integrated with the ICD and DSM. In addition, it can be used to: teach the PDM-2, for supervising diagnostic formulations, for progress reports and notes, as well as for research.

Limitations: The PDC2 is not a test in that it does not produce data independent of the clinician's insight. Rather it is a practitioner guide to organizing, charting, and contextualizing a complete PDM-2 formulation—a method for describing the patient along multiple domains to aid in diagnostic formulation and treatment planning.

References

Bornstein, R. F. (2011). From symptom to process: How the PDM alters goals and strategies in

- psychological assessment. *Journal of Personality Assessment*, 93, 142-150.
- Bornstein, R. F., & Gordon, R. M. (2012, Fall). What do practitioners want in a diagnostic taxonomy? Comparing the PDM with DSM and ICD. *Division/Review: A Quarterly Psychoanalytic Forum*, 6, 35.
- Gordon, R. M., Blake, A., Bornstein, R. F., Gazzillo, F., Etzi, J., Lingardi, V., McWilliams, N., Rothery, C., & Tasso, A.F. (2015) What do practitioners consider the most helpful personality taxa in understanding their patients? (manuscript submitted for publication).
- Gordon, R.M., & Bornstein, R.F. (2012). A practical tool to integrate and operationalize the PDM with the ICD or DSM. <http://www.mmipi-info.com/pdm-blog>.
- Gordon, R. M., Gazzillo, F., Blake, A., Bornstein, R. F., Etzi, J., Lingardi, V., McWilliams, N., Rothery, C., & Tasso, A. F. (in press). The relationship between theoretical orientation and countertransference awareness: Implications for ethical dilemmas and risk management. *Clinical Psychology and Psychotherapy*.
- Gordon, R. M., & Stoffey, R. W. (2014). Operationalizing the Psychodynamic Diagnostic Manual: A preliminary study of the Psychodiagnostic Chart (PDC), *Bulletin of the Menninger Clinic*, 78, 1-15.
- Gordon, R. M., Stoffey, R. W., & Perkins, B. L. (2013). Comparing the sensitivity of the MMPI-2 clinical scales and the MMPI-RC scales to clients rated as psychotic, borderline or neurotic on the Psychodiagnostic Chart. *Psychology: Special issue on Criminal Investigative Psychology*, 4, 9A, 12-16. doi: 10.4236/psych.2013.49A1003.
- Huprich, S., Lingardi, V., McWilliams, N., Bornstein, R. F., Gazzillo, F., & Gordon, R.M., (2015). The *Psychodynamic Diagnostic Manual (PDM)* and the *PDM-2*: Opportunities to significantly affect the profession. *Psychoanalytic Inquiry*, 35, 60-73.
- Lingardi, V., McWilliams, N., Bornstein, R. F., Gazzillo, F. & Gordon, R. M. (2015). The Psychodynamic Diagnostic Manual Version 2 (PDM-2): Assessing patients for improved clinical practice and research, *Psychoanalytic Psychology*, 32, 94-115.
<http://dx.doi.org/10.1037/a0038546>
- Weinryb, R. M., Rössel, R. J., & Asberg, M. (1991). The Karolinska Psychodynamic Profile. I. Validity and dimensionality. *Acta Psychiatrica Scandinavica*, 83(1), 64–72.

Zimmermann, J., Ehrental, J. C., Cierpka, M., Schauenburg, H., Doering, S., & Benecke, C. (2012). Assessing the level of structural integration using Operationalized Psychodynamic Diagnosis (OPD): Implications for DSM-5. *Journal of Personality Assessment*, 94(5), 522–532. doi:10.1080/00223891.2012.700664.

Psychodiagnostic Chart for Infancy and Early Childhood (PDC-IEC)

The PDC-IEC (Speranza, 2015) has been developed on the basis of the adult PDC as a user-friendly tool that would guide the clinician in the assessment of infant and young child as proposed by Infancy and Early Childhood Section of PDM-2.

The PDC-IEC requires the clinician to assess infant or child presenting problems and symptomatology, functional emotional developmental capacities; constitutional and maturational variations in the regulatory sensory processing profile of the child; relational patterns and disorders; and other medical or neurological diagnoses. Each of these dimensions helps not only to have an overall view of the child functioning, but has its value in the role it plays in pathogenesis.

Primary Diagnosis (Axis I) as well as other diagnoses are indicated and each diagnosis can be given a 1 = severe to 5 = mild rating.

Functional Emotional Developmental Capacities (Axis II) are rated along 6 expected emotional functions from 1 = severe deficits to 5 healthy.

Regulatory Sensory Processing Capacities (Axis III) are rated as regards sensory modulation, sensory discrimination and sensory-based motor functioning. An overall regulatory sensory profile is rated from 1 = severe defects to 5 = healthy.

Relational Patterns and Disorders (Axis IV) are rated on 8 dimensions for each significant caregiver, with an overall level of relational pattern and an attachment pattern evaluation.

Strengths. It guides the practitioner through a PDM-2 formulation for infant and young child using all relevant dimension of the assessment. It is user friendly and

provides a summary of the IEC Section. It can be used for clinical evaluation, treatment formulation, and outcome, as well as for empirical research.

Limitations. The PDC-IEC is not a test in the sense that it does not produce data independent of the clinician's judgment, nor are the scores compared to a standardized sample. Rather, it is a practitioner guide to charting a complete PDM-2 formulation for an infant or young child.

Psychodynamic Diagnostic Prototypes-2 (PDP-2)

The Psychodynamic Diagnostic Prototypes (PDP; Gazzillo, Lingiardi, & Del Corno, 2012) is a clinician-report instrument developed by Francesco Gazzillo and Vittorio Lingiardi, from the "Sapienza" University of Rome, together with Franco Del Corno, from the Association for the Research in Clinical Psychology in Milan. It is composed of 19 prototypic descriptions of personality disorders, one for each disorder of the P Axis of the PDM (first edition).

The PDP was developed to help clinicians and researchers complete the P Axis, even without a previous knowledge of the PDM. For this reason, Gazzillo et al. took the PDM descriptions of all the P Axis patterns/disorders, deleted the references presented in the manual, and simplified the PDM personality descriptions, which were more complex or inferential than needed in evaluating a patient. The descriptions have also been enriched by a careful study of the descriptions in other well-validated dynamic assessments.

In order to use the PDP, the clinician or rater assesses on a 5-point scale to what degree the patient resembles one or more PDP prototypes. A score of 1 means no resemblance while a score of 5 means a complete match between the patient's clinical presentation and the prototype. A categorical diagnosis of the disorder is indicated when the score is 4 or 5. Three to five sessions are usually required to provide the clinician with sufficient knowledge of the patient to make an accurate assessment.

The second edition of the PDP, the PDP-2 (Gazzillo, Genova & Lingiardi, 2015) is an adaptation of the PDP to the P Axis of the PDM-2. The basic structure of the instrument is

the same, but the PDP-2 takes into account the personality patterns described in the PDM-2 P Axis and assists the clinician also in the assessment of the level of personality organization of the patient and of the relevance of anaclitic or introjective themes in her/his personality.

For the assessment of the level of personality organization, the clinician or rater has to read a prototypical description of each level of personality organization (healthy, neurotic, borderline, psychotic) derived from the P Axis of the PDM-2 and then rate his or her patient on an 8-point Likert scale.

For identifying the prevalence of an anaclitic or introjective orientation, the clinician or rater takes into account both the personality pattern identified as more representative of the patient and a brief description of the main features of the two psychological orientations and then assesses the patient on two 5-point Likert scales.

Strengths. Most of the PDP personality patterns or disorders show good interrater reliability, concurrent and discriminant validity, and construct validity (Gazzillo, Lingardi, & Del Corno, 2012). The PDP descriptions are easy to use. Common emotional reactions to a patient with a given disorder are also described because these can assist in the assessment of the patient. This provides an educational value as well.

Limitations. Like all clinician-report assessments, the PDP-2 is only as accurate as a clinician's judgment, supported by clear descriptions of levels of personality organization and personality patterns. It was not possible to assess the convergent and discriminant validity of some of the PDP disorders because there were no empirical instruments assessing the same disorders as described by a different diagnostic manual.

References

Gazzillo, F., Lingardi, V., & Del Corno, F. (2012). "Toward the validation of three assessment instruments derived from the PDM2 P Axis: the Psychodynamic Diagnostic Prototypes, the Core Preoccupations Questionnaire and the Pathogenic

Beliefs Questionnaire". In *Bollettino di Psicologia Applicata*, 265, pp.31-45

Gazzillo, F., Lingardi, V., Del Corno, F., Genova, F., Bornstein, R., Gordon, R., McWilliams, N. (2015).

Clinicians' emotional responses and PDM2 personality disorders: a clinically relevant empirical study. *Psychotherapy*, 52 (2), 238-246.

The PDP-2 and its scoring manual are available from the Psychoanalytic Research Consortium website under "Publications", "related to PRC" (www.psychanalyticresearch.org).

B. Other Clinician-Rated Tools useful for PDM-2 Assessment

The following is a summary of other instruments available for clinician-based rating of core dimensions of personality functioning. See Table 1 for a synopsis of these assessment tools.

The Structured Interview for Personality Organization (STIPO)

Otto Kernberg and his group at the Personality Disorders Institute in White Plains, New York, developed the STIPO to operationalize the structural interview developed by Kernberg (1984). The STIPO comprises a structured interview of 97 questions that are used to assess the following domains of personality functioning: Identity, Object relations, Defenses, Reality testing, Coping and rigidity, Aggression, and Moral values. The patient's reply to each of the questions is scored on a 3-point Likert scale, from 0 (absence of pathology) to 3 (presence of pathology), with descriptive scoring anchors for each item. Moreover, the rater makes an overall estimation for each of the 7 domains on a 5-point scale from 1 (healthy) to 5 (pathological).

A group of researchers from Italy and America has developed an adolescent version of the STIPO, the Interview for Personality Organization Processes in Adolescence (IPOP-A), and its first validation data are promising (Ammaniti et al., 2014).

The STIPO and the IPOP-A may prove useful both for the assessment of the level of personality organization according to the P and PA Axis. Five of the domains are highly similar: identity, object relations, moral values/superego integration, reality testing, and coping and rigidity/resilience. The STIPO also has the domain of aggression, whereas the P and M-Axis assess more general emotional and relational capacities.

Strengths. The internal consistency of the two most important domains of the STIPO (identity and object relations) is high but is just acceptable for the reality testing domain and unknown for the other domains. The interrater reliability is acceptable, and its convergent and discriminant validity, compared to those values for other measures of similar constructs, is promising for the major domains of the STIPO.

Limitations. To achieve acceptable reliability, the STIPO and the IPOP-A require specific training, and a good knowledge of psychodynamic theory and of Kernberg's model of personality organization. It is also time-consuming (3-4 hours on average for the adult version and 2 hours for the adolescent version).

For more information, see www.personalitystudiesinstitute.com.

References

- Ammaniti, M., Fontana, A., Clarkin, A., Clarkin, J.F., Nicolais, G., Kernberg, O.F. (2014). Assessment of Adolescent Personality Disorders Through the Interview of Personality Organization Processes in Adolescence (IPOP-A): Clinical and Theoretical Implications. *Adolescent psychiatry*, 2, 1, pp.36-45.
- Stern, B.L., Caligor, E., Clarkin, J.F., Critchfield, K.L., Horz, S., MacCornack, V., Lenzenweger, M.F., Kernberg, O.F. (2010). Structured Interview of Personality Organization (STIPO): Preliminary Psychometrics in a Clinical Sample. *Journal of Personality Assessment*, 92, 1, 35-44.
- Hörz, S., Clarkin, J. F., Stern, B. L., & Caligor, E. (2012). The Structured Interview of Personality Organisation (STIPO): An instrument to assess severity and change of

personality pathology. In R. Levy, S. Ablon & H. Kächele (Eds.), *Psychodynamic Psychotherapy Research* (pp. 571-592). New York: Humana.

The Karolinska Psychodynamic Profile (KAPP)

The KAPP is a rating instrument founded upon psychoanalytic theory and developed in Sweden approximately 25 years ago by Weinryb and associates (1991a, 1991b) at the Karolinska Institute.

The clinician uses the material of a structured diagnostic interview such as the STIPO (see above) for the KAPP assessment. Clinician ratings are made on the basis of a detailed scoring system kept close to clinically observable phenomena. The scoring system generates 5 point scores (1; 1.5; 2; 2.5; 3) on 18 subscales.

The KAPP assesses relatively stable features of personality and mental functions, such as object relationships (intimacy and reciprocity, dependency and separation), ego strengths and flexibility (frustration tolerance, impulse control, capacity for regression in the service of the ego, coping with aggressive affects), body image and conceptions of body appearance and functioning, sexual life (sexual functioning and sexual satisfaction), social support (sense of belonging, sense of being needed, access to advice and help), alexithymic, normopathic, and controlling personality traits, and the overall level of personality organization.

Strengths. Just as with the STIPO, the KAPP may be useful for the assessment of the level of personality organization and its core functions (P Axis) and of several aspects of mental functioning (M Axis), such as the capacity of relationship and intimacy, impulse control and regulation, and adaptation, resiliency and strengths. The KAPP makes possible an assessment of personality and mental functioning from a psychoanalytic perspective that is psychometrically sound and that requires minimal training. KAPP scales show good validity and high reliability with minimal to modest training and can be used to identify personality features differentiating patients with different clinical disorders, to differentiate personality subtypes of patients with the same clinical disorder, and to assess the

personality changes of patients in the course of psychoanalysis and psychotherapy (Weinryb, Rossel, 1991a; Weinryb et al., 1991b; Turrina et al., 1996; Charitat, 1996; Weinryb et al., 1997; Wilczek et al., 1998; Weinryb, Gustavsson, & Barber, 2003).

Limitations. A reliable assessment of the KAPP requires familiarity with the underlying psychodynamic concepts. No empirically derived norms are yet available.

For more information, write to Alexander Wilczek (Alexander.Wilczek@ki.se).

References

- Charitat, H. (1996). *Evaluation de la personnalité: traduction et validation du Kapp*. Unpublished dissertation, French WorldChat
- Turrina C., Siani R., Regini C., Campana A., Bologna R. & Siciliani O. (1996). Inter-observer and test-retest reliability of the Italian version of the Karoliska Psychodynamic Profile [KAPP] in two groups of psychiatric patients. *Acta Psychiatrica Scandinavica* 93, 292-287.
- Weinryb, R.M., Rössel, R.J., Åsberg, M. (1991a). The Karolinska Psychodynamic Profile. I: Validity and dimensionality. *Acta Psychiatrica Scandinavica*, 83, 64-72.
- Weinryb, R.M., Rössel, R.J., Åsberg, M. (1991b). The Karolinska Psychodynamic Profile. II. Interdisciplinary and cross-cultural reliability. *Acta Psychiatrica Scandinavica*, 83, 73-76.
- Weinryb, R.M., Gustavsson, J.P., Barber, J.P. (2003). Personality Traits Predicting Long-term Adjustment after Surgery for Ulcerative Colitis. *Journal of Clinical Psychology*, 59, 9, 1015-1029.
- Wilczek, A., Barber, J.P., Gustavsson, J.P., Asberg, M., Weinryb, R.M. (2004). Change after Long-term Psychoanalytic Psychotherapy. *Journal of the American Psychoanalytic Association*, 52, 4, 1163-1184.

The Shedler-Westen Assessment Procedure (SWAP)

The Shedler Westen Assessment Procedure-200 (SWAP-200) (Shedler, 2015; Shedler & Westen, 1999a,b; Shedler, Westen & Lingardi, 2014) is a clinician-rated instrument that uses a Q-Sort method (explained below) to evaluate personality styles and disorders. The SWAP items were developed from a US nationwide survey of clinicians experienced in the treatment of personality disorders. After numerous iterations, a pool of 200 items was developed to organize clinical observations and inferences about a patient's personality and to provide an in-depth portrait of a patient's psychological functioning.

The SWAP-200 provides three score profile graphs that assess (a) DSM-IV and DSM-5 personality disorder diagnoses; (b) an alternative set of empirically-identified personality syndromes (called the Q factors) designed to be more clinically accurate and informative than DSM personality disorder diagnoses: schizoid, paranoid, antisocial/psychopathic, histrionic, narcissistic, obsessive, and dysphoric; and (c) twelve trait dimensions that highlight specific areas of personality functioning.

A revised version of the SWAP, the SWAP-II, was subsequently developed for research use (Westen & Shedler, 2007; Westen, Shedler, et al., 2012). SWAP-II differs from SWAP-200 in three respects. First, the SWAP-II has a slightly modified item set. Second, the newer edition possesses a different, empirically derived taxonomy of adult personality disorders subdivided in four spectra: (a) internalizing (depressive, anxious-avoidant, dependent-victimized, and schizoid-schizotypal disorders), (b) externalizing (antisocial-psychopathic, narcissistic, paranoid), and (c) borderline-dysregulated (borderline-dysregulated disorder), and (d) neurotic (obsessional and hysteric-histrionic disorder). Third, the SWAP-II was validated on a larger sample, more representative of the population ordinarily found in clinical practice, whereas the SWAP-200 validation sample contained only personality disordered patients.

An adolescent version of the SWAP, the Shedler-Westen Assessment Procedure for Adolescents (SWAP-II-A; Westen et al. 2005), follows a format similar to that of the adult SWAP. It uses a modified item set and an empirically derived taxonomy of adolescent

personalities articulated in two personality styles (healthy functioning and inhibited/self-critical) and five personality disorders (avoidant/constricted, antisocial/psychopathic, histrionic, narcissistic, and emotionally dysregulated). Reliability and validity are impressive (DeFife et al., 2013).

Using either version of the SWAP, the rater assigns a score from 0 (not descriptive) to 7 (most descriptive) to each of 200 items according to their descriptiveness of a patient's personality and functioning. The SWAP requires the clinician to use a fixed score distribution (i.e., assessors must assign each score a pre-specified number of times). Use of a fixed score distribution minimizes rater bias and maximizes reliability and validity (Block, 1978).

The SWAP items are most conveniently sorted using a web-based scoring program (available at www.SWAPassessment.org). Users can select three diagnostic reports, including a Clinical Interpretive Report that includes a comprehensive interpretation of all test scores, as well as personality diagnoses, clinical case formulations, and treatment recommendations.

Other than a description of pathological personality features, the SWAP also provides descriptors of positive mental health, represented in twenty-four of the 200 items, and generates a profile of a healthy, high-functioning individual.

The 200 items of the SWAP, and the 7 degrees of salience assigned to each item, generate nearly innumerable permutations that permit the capture of complex personality patterns (Shedler 2015). Further studies (Lingiardi, Shedler & Gazzillo, 2006; Waldron et al. 2011) have shown that the 30 items receiving the highest scores provide a useful summary of patient functioning, such that they can be used as the core of a case formulation.

To improve the SWAP's utility for evaluating a patient's level of health-sickness, and the patient's areas of health and sickness, two supplemental score indices were developed by a group of researchers in New York and have been described, along with evidence for their validity (Waldron et al., 2011).

The first index, the Personality Health Index (PHI), provides a global assessment of

personality health using the 200 SWAP item scores. The index is based upon pooled clinical judgments of the health significance of each SWAP item, multiplied by the actual score (0 to 7) given to the patient by the rating clinician. This total raw score is then automatically converted into a percentile score in reference to a national sample of 307 SWAPs from different points in the treatments of 70 psychoanalytic patients (Cogan & Porcerelli, 2004, 2005). The percentile standing provides a ready and understandable metric of overall psychological health or illness.

The second index derived from the SWAP, the RADIO, delineates an individual's particular strengths and difficulties within five central domains of personality functioning: Reality testing, Affect regulation and tolerance, Defensive organization, Identity integration, and Object relations (with the acronym RADIO). It also displays both healthy and pathological items in these areas in five separate charts. The modified SWAP software also charts these scores in comparison with other points in the treatment, or with other patients, as desired by the user, so it is possible to follow a patient's progress over the course of a treatment. These domains are similar to those used by the PDM-2 to describe Levels of Personality Functioning.

Two published case examples demonstrate how this may prove useful clinically. While these indices have obvious applications for psychotherapeutic and psychoanalytic research, they comprise a useful resource for clinical work and training as well (Waldron et al., 2013; Gazzillo et al. 2014).

The SWAP, in its different versions, can be a useful support for P and PA Axes assessment, while the PHI and RADIO categories may help the clinician in the assessment of the patient on the M and MA Axes.

Strengths. The SWAP provides systematic personality diagnoses and case formulations by allowing clinicians to record their clinical observations systematically and then applying statistical and psychometric methods to the resulting data to optimize diagnosis and clinical case description. The information is useful to treating clinicians in developing clinically rich case formulations that can guide effective intervention. It shows

excellent validity and reliability (Shedler, 2015). Clinicians and researchers can use the SWAP-200 to compare a patient's overall health-sickness at different points in treatment, providing a valuable method for demonstrating change over the course of psychotherapy (Waldron et al. 2013, Gazzillo et al. 2014). The items are straightforward, jargon-free clinical statements about the individual being rated. Its reliability and validity have been reported and extended over the past 10 years (Shedler, 2015; Westen & Muderrisoglu, 2006; Shedler & Westen, 2006).

Limitations. The SWAP requires between 25 and 40 minutes to score (depending on the assessor's familiarity with the instrument), an amount of time not readily available in many clinical settings. Moreover, for a reliable SWAP assessment, the clinician or rater should know the patient well; the SWAP authors recommend a minimum of six clinical sessions for patients in treatment or, for individuals being seen for assessment only, a systematic interview for personality assessment, the Clinical Diagnostic Interview (CDI), which can be administered in approximately 2½ hours

For more information, visit www.swapassessment.com.

References

- Blagys, P.S., Bi, W., Shedler, J., & Westen, D. (2012). The Shedler-Westen Assessment Procedure (SWAP): Evaluating psychometric questions about its reliability, validity, and impact of its fixed score distribution. *Assessment*, 19, 3, 370-382.
- DeFife, Jared A; Malone, Johanna C; DiLallo, John; Westen, Drew. (2013). Assessing adolescent personality disorders with the Shedler-Westen Assessment Procedure for Adolescents. *Clinical Psychology: Science and Practice*, 20: 393-407.
- Gazzillo, F, Waldron, S., Genova, F., Angeloni, F., Ristucci, C. and Lingiard, V. (2014). An empirical investigation of analytic process: contrasting a good and poor outcome case. *Psychotherapy*: 270-282.
- Shedler, J. (2015). Integrating clinical and empirical perspectives on personality: The Shedler-Westen Assessment Procedure (SWAP). In SK Huprich, (Ed.), *Personality*

Disorders: Toward Theoretical and Empirical Integration in Diagnosis and Assessment. Washington, D.C.: American Psychological Association.

Shedler, J. (2000). A new language for psychoanalytic diagnosis.

Psychologist-Psychoanalyst, 20, 30-37.

Shedler, J., Westen, D., Lingiardi, V. (2014). *La valutazione della personalità con la SWAP-200.* Raffaello Cortina, Milano.

Waldron, S., Gazzillo, F., Genova, F., & Lingiardi, V. (2013). Relational and classical elements in psychoanalyses: An empirical study with case illustrations. *Psychoanalytic Psychology*, 30, 567–600.

Waldron, S., Moscovitz, S., Lundin, J., Helm, F.L., Jemerin, J., Gorman, B, (2011).

Evaluating the outcome of psychotherapies: the Personality Health Index.

Psychoanalytic Psychology, 28: .363-388.

Westen, Drew; DeFife, Jared A; Malone, Johanna C; DiLallo, John. (2014). An empirically derived classification of adolescent personality disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*. 53: 528-549.

Westen, Drew; Dutra, Lissa; Shedler, Jonathan. (2005). Assessing adolescent personality pathology. *British Journal of Psychiatry*. 186. 227-238.

Westen, D., & Shedler, J. (1999a). Revising and assessing Axis II, Part 1: Developing a clinically and empirically valid assessment method. *American Journal of Psychiatry*, 156, 258-272.

Westen, D., & Shedler, J. (1999b). Revising and assessing Axis II, Part 2: Toward an empirically based and clinically useful classification of personality disorders. *American Journal of Psychiatry*, 156, 273-285.

Westen, D., Shedler, J., Durrett, C., Glass, S., & Martens, A. (2003). Personality diagnoses in adolescence: DSM-IV Axis II diagnoses and an empirically derived alternative. *American Journal of Psychiatry*, 160, 952-966.

Westen, D., Shedler, J., Bradley, B., & DeFife, J.A. (2012). An empirically derived taxonomy for personality diagnosis: Bridging science and practice in conceptualizing personality. *American Journal of Psychiatry*, 169(3), 273-284.

The Operationalized Psychodynamic Diagnosis (OPD)

The OPD Task Force was founded in 1992 in Germany by a group of psychoanalysts, specialists in psychosomatic medicine, and psychiatrists with a background in psychotherapy research. The goal of the group was to broaden the descriptive and symptom-oriented ICD-10 classification system to include fundamental psychodynamic dimensions. This task force developed a diagnostic inventory and a handbook designed for training and clinical purposes for experienced therapists. Since then, the manual has been translated into 11 languages. Further translations are in preparation.

In 1996 the OPD Task Force published its first manual and in 2006 the second version, OPD-2, was published in Germany (OPD Task Force, 2006; English, 2008). The OPD-2 additionally includes a textbook that manualizes the focus selection for individual treatment planning. On the basis of therapeutic foci, it is possible to evaluate the treatment outcome and follow-up. Additionally, the Heidelberg Structure Change Scale allows the evaluation of clinical changes (Grande et al., 2003).

Recently, the second version of the German-speaking manual of the OPD for child and adolescent psychotherapists was published (OPD Taskforce, 2014). A self-report instrument to measure the patient's level of structural integration is also available (Ehrenthal et al., 2012). Additional resources, such as modules for substance abuse, forensic patients and trauma, are under development.

The OPD requires the clinician to complete 5 Axes:

Axis I: Experience of illness and prerequisites for treatment. The clinician characterizes the severity of illness, degree of suffering, subjective illness attributions (the patient's understanding of the cause of his or her illness), models of change (the patient's ideas about how treatment may help), and resources (personal resources such as ego strengths, humor, intellect, and psychosocial support by family, friends, companions).

Axis II: Interpersonal Relations. The clinician characterizes the dysfunctional habitual relationship experiences of the patient and describes the patient's relationship to the therapist as well.

Axis III: *Conflict* The clinician describes the patient according to seven inner conflicts: individuation vs. dependency; dependence vs. self-sufficiency; control vs. submission; valuing self vs. valuing others; guilt conflict; oedipal conflict; and identity conflict.

Axis IV: *Structure*. The clinician rates 8 dimensions, each with 3 facets: (a) self-perception (related to the patient's capacity for self-reflection, and identity); (b) object perception (related to the patient's capacity to differentiate self from object); (c) self-regulation (e.g. impulse control); (d) regulation of object relationships and interests; (e) internal communication (fantasies, body awareness); (f) communication with others (being in contact, empathy), (g) attachment capacity to internal objects (use of introjects, variability of introjects), and (h) attachment capacity to others (including the capacity to separate). Ratings are combined in a global score.

Axis V: *Mental disorders, personality disorders, and psychosomatic disorders* (in conjunction with chapter V (F) of the ICD 10 or DSM-IV/ DSM-5).

After an initial one- to two-hour interview, the clinician (or a clinical observer) evaluates the patient's psychodynamic profile on the basis of the OPD-2 Axes and categories. The OPD interview technique alternates between relatively unstructured phases of free exploration and more structured questions. The interviews are designed to elicit all information that is relevant for assessing an interviewee's maladaptive relationship patterns, enduring motivational conflicts, and structural capabilities. Guidelines on how to conduct the interview are sufficiently flexibly so that the interview can still be conducted as an open psychodynamic interview.

The OPD can be applied by experienced raters with relatively good reliability and can be recommended for teaching. The reliability has been found to be higher for the structural Axis when it is rated from videotaped interviews by experienced raters than by inexperienced raters (Doering et al. 2014, Benecke et al., 2011; for summary see Zimmermann et al. 2012).

Strengths: The OPD can be a very useful aid for the assessment of Axes P, M, S, PA, MA and SA.

Trained and experienced clinicians can reliably apply the OPD rating to clinical interview material (Cierpka et al., 2007) and its validity is well supported in 17 samples, including data from more than 2,000 participants (Zimmermann et al., 2012). An OPD expert consensus study confirmed that the OPD structure Axis is highly similar to the proposed Level of Personality Functioning Scale (LPFS) of the DSM-5, Personality Disorders Section (Zimmermann et al., 2012). Both measures share a common severity dimension. Most important, the global score of OPD structure rating shows substantial and specific associations with the presence of DSM-assessed personality disorders, associations supporting its use as an index for the severity of personality pathology. In a second study, the findings also indicate that DSM-5 level of personality functioning and OPD level of structural integration are highly interrelated constructs. Zimmermann et al. (2013) demonstrate that this conceptual overlap is empirically reflected in high correlations between the independent OPD structure Axis and LPFS ratings that are based on the same video material.

The OPD fulfills Strupp's recommendation of the need for "problem-treatment-outcome congruence" in psychodynamic psychotherapy (Strupp & Schacht, 1988). The system is also applicable for pre-post-follow up evaluations of psychoanalytic therapies (Rudolf et al., 2002). Besides its use as a research-oriented diagnostic tool and as a tool to evaluate changes in patient's mental functioning, the OPD most importantly targets the field of therapeutic activity by defining the clinically most relevant therapy foci for patients in outpatient and inpatient settings. The use of the OPD in therapy planning facilitates evaluation of relationships, conflict, and structure to decide whether the treatments should be conducted in a more conflict-oriented or more structure-oriented way. By providing a standardized format, it helps clinicians to generate a consistent psychodynamic formulation of intrapsychic and interpersonal dynamics.

Limitations: The OPD system is a comprehensive approach. To learn the basic use of this tool requires at least 60 hours of training. Completion of the evaluation sheets of all the Axes takes approximately 30 minutes. So far, the OPD is limited only to therapy

planning, with no attempt to manualize intervention processes. However, Operationalized Psychodynamic Therapies (OPT) for different disorders are currently being developed by the OPD group.

For more information, see www.opd-online.net.

References

- Arbeitskreis OPD (Hrsg.). (1996). *Operationalisierte Psychodynamische Diagnostik. Grundlagen und Manual*. Bern: Huber.
- Arbeitskreis OPD (Hrsg.). (2006). *Operationalisierte Psychodynamische Diagnostik OPD-2. Das Manual für Diagnostik und Therapieplanung*. Bern: Huber.
- Arbeitskreis OPD-KJ. (Hrsg.). (2014). *Operationalisierte Psychodynamische Diagnostik im Kindes- und Jugendalter (OPD-KJ-2). Grundlagen und Manual*. Bern: Huber.
- Benecke, C., Bock, A., Wieser, E. et al. (2011) Reliabilität und Validität der OPD-KJ-Achsen Struktur und Konflikt. *Prax. Kinderpsychol. Kinderpsychiat.* 60, 60-73.
- Cierpka, M., Rudolf, G., Grande, T., & Stasch, M. (2007). Operationalized Psychodynamic Diagnostics (OPD). Clinical relevance, reliability and validity. *Psychopathology*, 40, 209-220.
- Doering, S., Burgmer, M., Heuft, G., Menke, D., Bäumer, B., Lübking, M., Feldmann M., Schneider, G. (2014) Diagnosing Personality Functioning: Validity of the Operationalized Psychodynamic Diagnosis (OPD-2) Axis IV: Structure. *Psychopathology*, 47, 185-93.
- Ehrenthal, J.C., Dinger, U., Horsch, L., Komo-Lang, M., Klinkerfuß, M., Grande, T & Schauenburg, H. Der OPD-Strukturfragebogen (OPD-SF): Erste Ergebnisse zu Reliabilität und Validität. *Psychother Psych Med.* 2012;62:25-32.
- Grande, T., Rudolf, G., Oberbracht, C., & Pauli-Magnus, C. (2003). Progressive changes in patients' lives after psychotherapy: which treatment effects support them? *Psychotherapy Research*, 13(1), 43-58.

OPD Task Force (Ed.). (2008). *Operationalized Psychodynamic Diagnostics OPD-2. Manual of Diagnostics and treatment planning*. Göttingen: Hogrefe & Huber.

Rudolph, G., Grande, T., Dilg, R., Jakobsen, T., Keller, W., Oberbracht, C. et al. (2002). Structural changes in psychoanalytic therapies – The Heidelberg-Berlin study on long-term psychoanalytic therapies (PAL). In M. Leuzinger-Bohleber & M. Target (eds.), *Outcomes of psychoanalytic treatment – perspectives for therapists and researchers* (pp. 201-222), London, Whurr.

The Psychodynamic Functioning Scales (PFS)

Per Høglend and his colleagues primarily in Norway have carried out extensive systematic studies of processes and outcomes of dynamic psychotherapies. Recognizing the need for better outcome studies, they started more than 25 years ago to develop instruments to respond to this imperative (Høglend, 1995; Høglend et al., 2000).

The PFS uses 6 scales, with the same format, 0-100, as the GAF to measure psychological functioning over the three previous months. Three of the scales measure interpersonal aspects: Quality of Family Relationships, Quality of Friendships, and Quality of Romantic/Sexual Relationships. The other three measure intrapersonal functioning: Tolerance for Affects, Insight, and Problem Solving Capacity. The evaluation is based on interview data (1-2 hours). The scales have been translated to German, French, and Portuguese.

Strengths. The PFS can be a useful aid in the assessment of M Axis functions (relationship and intimacy, affective range, communication and understanding, psychological mindedness) and overall mental functioning.

Aspects of content validity, internal domain construct validity, discriminant validity from symptom measures, and sensitivity for change in dynamic therapy have been established in different samples of patients and evaluators (Bøgwald & Dahlbender, 2004; Hagtvet & Høglend, 2008; Hersoug, 2004; Høglend et al., 2000; Høglend, 2004). The instrument is brief, and there are many levels, to increase sensitivity for change.

Limitations. So far, reliability has only been tested using clinicians with several years of dynamic training.

For more information, write to Per Høglend (p.a.hoglend@medisin.uio.no).

References

- Høglend P. Dynamic Scales. Manual. Oslo: Dep of Psychiatry, University of Oslo 1995
- Høglend P, Bøgwald KP, Amlo S, Heyerdahl O, Sørbye Ø, Marble A, Sjaastad MC, Bentsen H. Assessment of change in dynamic psychotherapy. *J Psychother Pract Res* 2000;9:190-199
- Hagtvet, K A Høglend, P. Assessing precision of change scores in psychodynamic psychotherapy: A generalizability theory approach. *Measurement and Evaluation in Counseling and Development*, 2008; 41: 162-177
- Høglend P, Bøgwald KP, Amlo S, Marble A, Sjaastad MC, Sørbye Ø, Heyerdahl O, Johansson P, Ulberg R. Transference interpretations in dynamic psychotherapy: Do they really yield sustained effects? *Am J Psychiatry* 2008; 165; 763–771
- Høglend, P. Hersoug, A.G., Bøgwald, K.P., Amlo, S., Marble, A., Sørbye, Ø., Røssberg, J.I., Ulberg, R., Gabbard, G. & Crits-Christoph, P. (2011). Effects of Transference Work in the Context of Therapeutic Alliance and Quality of Object Relations. *Journal of Consulting and Clinical Psychology* 79, 697–706.

The Scales for Psychological Capacities (SPC)

The SPC is an interview-based instrument developed by Robert Wallerstein and colleagues at the University of California, San Francisco, in the 1980s. The Scales were designed to profile personal functioning that can reflect treatment-related changes in underlying personality organization.

The SPC comprises 36 Likert-type scales, measuring 17 dimensions of positive functioning or "capacities," that are applied by trained raters to semistructured clinical interviews or recorded treatment sessions. The healthy end of each of the 17 dimensions is

anchored by a title and definition. The dimensions are self-coherence, self-esteem, zest for life, hope, flexibility, responsibility, persistence, standards and values, affect regulation, impulse regulation, regulation of sexuality, assertion, empathy, trust, reliance on others, relationship commitment, and reciprocity. Each of these dimensions has between one and three unhealthy opposites (e.g., self-esteem vs. grandiosity and self-depreciation), yielding 36 subdimensions. Each subdimension is rated on a 4-point Likert scale from fully adaptive functioning (0) to seriously or obviously compromised (3). Definitions and clinical examples anchor each scale point.

Wallerstein enlisted the help of clinicians from many countries and theoretical orientations in developing the capacity set. They define psychological capacities as psychological resources needed to achieve adaptive functioning and life satisfaction (DeWitt et al., 1991). The scales are designed to be theoretically informed but not theory specific so that the scales can contribute to evaluating the effects of various treatments on personality and functioning. The SPC is written in plain English in order to minimize ambiguity and optimize its ease of use and ability to be translated. An adolescent version is available.

Strengths. The SPC may be useful as a support in the assessment of a patient according to the M and MA Axis capacities (identity, quality of internal experience, resiliency and relational and intimacy capacity, in particular). The instrument has demonstrated good interrater and test-retest reliability in studies by groups in California (De Witt et al., 1999) and in Germany (Leuzinger-Bohleber & Fischmann 2007; Huber et al., 2004, 2005) and also has been found to have content and construct validity.

The SPC taps into areas that are important to an individual's general, positive functioning. Changes in these areas go beyond the symptom changes, and impaired functioning on these scales reflects the typical kinds of impairments seen in psychiatric outpatients. Substantial positive changes in the Munich Psychotherapy Study in the patients' psychological capacities during psychoanalytic treatment demonstrate the measure's utility (Huber et al., 2005).

Limitations. Training requires four to six hours to achieve reliability. The assessment tool has not yet been applied to sufficient numbers of patients early and late in treatment to generate findings that would fully reveal the potential (or weaknesses) of the instrument.

For more information, write to Kathryn DeWitt (kndewitt@gmail.com).

References

- DeWitt, K. N., Hartley, D. E., Rosenberg, S. E., Zilberg, N. J., & Wallerstein, R. S. (1991). Scales of Psychological capacities: Development of an assessment approach. *Psychoanalysis and Contemporary Thought*, 14, 343-361.
- DeWitt, K. N., Milbrath, C., & Wallerstein, R.S. (1999). Scales of psychological capacities: Support for a measure of structural change. *Psychoanalysis and Contemporary Thought*, 22, 453-80.
- Huber, D., Brandl, T., & Klug, G. (2004). The Scales of Psychological Capacities (SPC): Measuring beyond symptoms. *Psychotherapy Research*, 14, 89-106.
- Huber, D., Henrich, G., & Klug, G. (2005). The Scales of Psychological Capacities: Measuring change in psychic structure. *Psychotherapy Research*, 15, 445-456.
- Huber, D., Henrich, G., & Klug, G. (2013). Moderators of Change in Psychoanalytic, Psychodynamic And Cognitive Behavioral Therapy, *Journal of the American Psychoanalytic Association* , 61(3), 585-589.
- Leuzinger-Bohleber, M., & Fischmann, T. (2007). Application of the scales of psychological capacities in a multiperspective, representative follow-up. In W. Bucci & N. Freedman (Eds.), *From Impression to Inquiry: a Tribute to the Work of Robert Wallerstein* (pp. 82-96). London: The International Psychoanalytic Association.
- Zilberg, N. J., Wallerstein, R. S., DeWitt, K. N., Hartley, D. E., & Rosenberg, S. E. (1991). A Conceptual analysis and strategy for assessing structural change. *Psychoanalysis and Contemporary Thought*, 14, 317-342.

The Social Cognition and Object Relations Scale (SCORS)

The SCORS was originally created by Westen (1995) to systematically assess (via free-response data) multiple dimensions of cognitive and affective processes that mediate interpersonal functioning beyond the overt presentation of the patient. It integrates clinically based object relational theories with experimentally generated theories of social cognition (Stein et al., 2012). There are three versions of the SCORS. The focus on this section will be on the SCORS-Global Rating Method (SCORS-G).

The SCORS-G consists of eight variables that are scored on 7-point anchored scales on which lower scores (i.e., 1, 2, or 3) indicate more pathological responses and higher scores (i.e., 5, 6, or 7) indicate healthier and more mature aspects of object relations. The eight variables are as follows:

1. Complexity of Representations of People (COM), which assesses presence, degree, and differentiation of internal states and relational boundaries;
2. Affective Quality of Representations (AFF), which examines the emotional lens through which a person views his or her environment;
3. Emotional Investment in Relationships (EIR), which assesses the level of intimacy and emotional sharing;
4. Emotional Investment in Moral Standards (EIM), which measures how a person views others and acts in relation to morality and compassion for others;
5. Understanding of Social Causality (SC), which evaluates the extent to which the person understands human behavior as well as the narrative's coherence, logic, and reasoning;
6. Experience and Management of Aggressive Impulses (AGG), which explores the person's ability to tolerate and manage aggression;
7. Self Esteem (SE)
8. Identity and Coherence of Self, (ICS), which is the degree to which a person has an integrated sense of who he or she is.

The SCORS-G can be applied to numerous forms of narrative data, such as

psychotherapy sessions, Thematic Apperception Test (TAT) data, early memory narratives, and Relationship Anecdote Paradigm (RAP) and clinical interviews.

The SCORS-G has been able to successfully discriminate between aspects of psychopathology and personality (Ackerman, Clemence, Weatherill, & Hilsenroth, 1999; Stein et al., 2012). Other studies have explored the SCORS-G as it relates to attachment and trauma (Calabrese, Farber, & Westen, 2005; Ortigo, Westen, DeFife, & Bradley, 2013), as well as to other clinical phenomena (Stein et al., in press). In addition, there have been numerous studies that have used the SCORS-G in the context of process and outcome research in psychotherapy (Fowler, Ackerman, Speanburg, Bailey, Blagys, & Conklin, 2004).

For those who do not want to rate narrative data formally (in the context of research), a clinician can use variable or domain anchor points as an alternative (1, 3, 5, 7). Using the SCORS-G anchor points in a clinician's clinical work can help him or her describe, explain, or explore relevant dimensions of patients' mental functioning and the impact of the therapeutic process.

Strengths. The SCORS-G can be a very useful aid for the assessment of several capacities of the M and MA Axis, such as the capacities for differentiation and integration, the quality of internal experience, the capacity for relationship and intimacy, and the capacity to construct internal standards and values.

It is one of the most common psychodynamic measures used to rate narrative data. Anchor points are worded in an experience-near way useful to clinicians from various orientations. Interrater reliability and convergent and discriminant validity have been consistently demonstrated in previous research (Stein et al., 2013).

Limitations. The most significant weakness of the SCORS-G is its limited normative data. Also, there are medium to high intercorrelations between the eight variables (Stein et al., 2012). Other weaknesses include the amount of time it takes to reach good reliability in its assessment. However, there are training manuals (Stein et al., 2011).

SCORS-G ratings may vary across types of narratives, making direct comparisons more challenging. Lastly, some of the SCORS-G variables have default ratings because not every dimension is routinely present in narratives. The default ratings are often (but not exclusively) given a middle rating of 4. Higher ratings may then be more reflective of absence of data for that variable than of more mature aspects of object relations.

For more information, write to Michelle Stein (steinmichelle@yahoo.com).

References

- Ackerman, S., Clemence, A., Weatherill, R., & Hilsenroth, M. (1999). Use of the TAT in the Assessment of DSM-IV Cluster B Personality Disorders. *Journal of Personality Assessment*, 73, 422-448.
- Stein, M., Hilsenroth, M., Slavin-Mulford, J., & Pinsker, J. (2011). Social Cognition and Object Relations Scale: Global Rating Method (SCORS-G; 4th ed.). Unpublished manuscript, Massachusetts General Hospital and Harvard Medical School, Boston, MA
- Stein, M.B., Slavin-Mulford, J., Sinclair, S.J., Siefert, C.J., & Blais, M.A. (2012). Exploring the Construct Validity of the Social Cognition and Object Relations Scale in a Clinical Sample. *Journal of Personality Assessment*, 94(5), 533-540.
- Westen, D. (1995); "Social Cognition and Object Relations Scale: Q-sort for Projective Stories (SCORS-Q);" Unpublished manuscript; Department of Psychiatry, The Cambridge Hospital and Harvard Medical School, Cambridge, MA.

The Psychodynamic Conflict Rating Scales (PCRS)

The first version of the PCRS, developed by J. Christopher Perry starting in 1980, derived from the tradition of understanding intrapsychic or psychological conflict, then broadened to include recent developments in both self psychology and object relationship perspectives. Each conflict consists of a series of statements that reflect some aspect of conflict among motives (wishes and fears), internal states, external events (especially

interpersonal interactions), and defensive functioning. The result of conflict may lead to either symptom formation or compromise formations, which are also noted. However, because conflict is internal and partly or wholly unconscious, we can only infer conflict indirectly through what can be observed as anomalies or disturbances.

The PCRS assess 14 defined conflicts, divided into 7 Focal conflicts, which are presumed to be developmentally later in origin and to be largely synonymous with oedipal level conflicts, and 7 Global conflicts, presumed to have earlier developmental origins and to be largely synonymous with oreoedipal level conflicts.

The Focal conflicts include (1) Dominant Other, (2) Dominant Goal, (3) Counter-dependent, (4) Ingratiation-disappointment, (5) Ambition-Achievement, (6) Competition-hostility, and (7) Sexual Pleasure vs. Guilt.

The Global conflicts include (8) Overall Gratification Inhibition, (9) Separation-Abandonment, (10) The Global Conflict over the Experience and Expression of Emotional Needs and Anger, (11) Object-hunger, (12) Fear of Fusion, (13) Rejection of Others, and (14) Resentment over Being Thwarted by Others.

The PCRS also has 14 companion scales that assess healthy adaptation to each conflict, so that both pathology and healthy adaptation are represented separately (Perry, 1997, 2006; Perry & Cooper, 1986).

For clinical use, a thorough history in the context of dynamic interviewing or therapy sessions is sufficient. For research purposes, the author suggests two data sources: (a) a dynamic intake or follow-up type interview, and (b) an interview that focuses on relationship stories, such as the Relationship Anecdote Paradigm (RAP) interview (Beck & Perry, 2008).

The original version of the PCRS consisted of a definition of each conflict and a section on discriminating it from near-neighbor conflicts. The rater then filled out a 4-point pattern identification score that is readily used. The current version improved upon this psychometrically by assessing each conflict with a series of 8 to 15 items, each of which is a low-inferential statement of some aspect of the conflict, scored as (0) absent, (1) somewhat true, or (2) definitely present or true. Statements represent affective, behavioral,

or cognitive facets of the conflict, and are generally in nontechnical English. Scores for each conflict are converted to reflect the proportion of the maximum score (scaled from 0 to 1.00).

Interrater reliability determined on an early version was adequate to good (Perry & Cooper, 1986; Perry & Perry, 2004) while more recent studies generally find internal consistency and interrater reliability to be good to excellent. (Perry & Constantinides, manuscript under review). Cutoffs define the definite presence or absence of a conflict or the presence of healthy adaptation. Additional separate scales summarize the (a) affective, (b) behavioral, and (c) cognitive features of each conflict. Mean summary scores are then calculated for the Focal, Global, Overall Pathological Conflict, and Adaptation to Conflict scales. Two studies examining change with short-term or longer term treatments have demonstrated convergent and discriminant validity (Perry & Constantinides, manuscript under review).

Strengths. The PCRS can give useful data for the assessment of Axes P and PA and for assessing the capacity of relationships and intimacy on the Axes M and MA. Clinicians could use the original pattern-recognition version with minimal time required for rating. The conflicts are intuitively understandable, and the statements readily reflect clinical observation, as well as conflict issues to which clinicians attend. The method has demonstrated usefulness for measuring change over time and can determine when individuals are dynamically recovered.

Limitations. Training requires rating about three to five training cases and intermittent calibration by doing consensus ratings. Although the PCRS could be rated after a single dynamic interview, the authors prefer several interviews to ensure adequate coverage of data. Research use requires recorded interviews. Ratings require about 90-120 minutes per session rated.

For more information, write to John Christopher Perry (jcristopher.perry@mcgill.ca).

References

- Beck, S., & Perry, J.C. (2008). The measurement of interview structure in five types of psychiatric and psychotherapeutic interviews. *Psychiatry: Interpersonal and Biological Processes* 71, 219-233.
- Perry JC (1997). The Idiographic Conflict Formulation Method. *Handbook of Psychotherapy Case Formulation*. Eells, T, ed., Guilford Press, New York, 1997, chapter 5, pp. 137-165.
- Perry, J. (1994). Assessing Psychodynamic Patterns Using the Idiographic Conflict Formulation Method. *Psychotherapy Research*, 4, 239-252.
- Perry, J.C. (2006). *The Psychodynamic Conflict Rating Scales (PCRS)*, 3rd edition, 2006. Published by the author, Available from the author.
- Perry, J.C., & Constantinides, P., (manuscript under review). Dynamic conflicts in recurrent major depression: does combined short-term psychotherapy and medications lead to healthy dynamic functioning?
- Perry JC, & Cooper SH. (1986). A Preliminary Report on Defenses and Conflicts Associated with Borderline Personality Disorder. *Journal of the American Psychoanalytic Association* 1986; 34:865-895.
- Perry, J.D., & Perry, J.C. (2004). Conflicts, defenses and the stability of narcissistic personality features. *Psychiatry*, 67, 310-330.

The Object Relations Inventory (ORI)

The ORI is an unstructured method developed by Blatt and colleagues for assessing a person's representations of self and significant figures, that is, a person's object world. It derives originally from work on the use of parental descriptions as a method for assessing object representations (Blatt, Wein, Chevron, & Quinlan, 1979), on the basis of cognitive developmental theory and developmental object relations theories current in psychoanalysis in the 1970s (e.g., Fraiberg, A. Freud, Jacobson, Kernberg, Kohut, Mahler, Winnicott). Since the 1980s, with the influence of developmental theorists like Stern, and Beebe and Lachmann, as well as of intersubjectivity theory rooted in Hegelian philosophy

(e.g., Aron, Benjamin, Fonagy and colleagues), the ORI has come to use the individual's descriptions of his or her significant figures to assess the ability to understand both oneself and one's intersubjective interpersonal matrix.

The ORI's current form, as an interview, was constructed by Sugarman (personal communication to J. Auerbach, August 29, 2014). Subjects are asked to describe parents, significant others (including pets), self, and therapist, with inquiry by the interviewer.

The content and cognitive structural organization of the descriptions are evaluated, and several methods have been developed. The best known of these methods is the Conceptual Level scale (CL; Blatt, Chevron, Quinlan, Schaffer, & Wein, 1988), which describes a developmental progression of five levels of the object: sensorimotor, perceptual, external iconic, internal iconic, and conceptual. In addition, Blatt and colleagues (Blatt et al. 1979; 1988; Quinlan, Blatt, Chevron, & Wein, 1992) evaluated the descriptions for three qualitative or thematic factors: Benevolence, Punitiveness, and Striving. Also scored is the degree of ambivalence expressed when describing the figure. Interrater reliability for CL and the three qualitative-thematic factors were found to be adequate ($r > .77$). Later factor-analytic studies (e.g., Heck & Pincus, 2001; Huprich et al., in press) developed factors of Agency, Communion, and Punitiveness.

Subsequently Blatt's research group (Diamond, Blatt, Stayner, & Kaslow, 1991; Diamond, Kaslow, Coonerty, & Blatt, 1990) constructed the Differentiation-Relatedness (D-R) Scale, with 10 levels, as follows: boundary disturbance (Levels 1 and 2); mirroring, unilateral idealization or denigration, and polarization (Levels 3, 4, and 5), emergent and consolidated self and object constancy (Levels 6 and 7), and higher transformations of empathy, mutuality, and intersubjectivity (Levels 8, 9, and 10). A D-R score of 6 to 7 (i.e., reflective of some degree of object constancy) is considered a threshold for normal functioning. Interrater reliability for the new scale was found to be adequate (intraclass correlation coefficient = .83).

Blatt and colleagues (Blatt, Bers, & Schaffer (1992) also developed a rating manual to capture the structural complexities involved in the description of self. Eighteen reliable dimensions, including CL, were resolved into five factors: Agency, Reflectivity,

Differentiation, Relatedness, and Relatedness to the Examiner.

Reviews of research findings with the various ORI methods (e.g., Blatt, 2008; Blatt & Auerbach, 2001; Blatt, Auerbach, & Levy, 1997; Priel, 2005, Bender et al., 2011; Huprich, Auerbach, Porcerelli, & Bupp, in press) show that structural aspects of object relations (e.g., CL, D-R) are moderately consistent across individuals described by a given subject (Besser & Blatt, 2007; Lowyck, Luyten, Verhaest, Vandeneede, & Vermote, 2013; Priel, 2005). Variables like CL and D-R are thus considered to be structural dimensions, reflective of the person's underlying level of organization, regardless of whether significant figures are described in predominantly positive or predominantly negative terms, and a significant figure description can be very negative in content and yet still organized at a high developmental level. Two studies of long-term inpatient treatments of patients with borderline personality organization have found that such patients usually begin treatment with D-R scores indicative of polarization and splitting (D-R 5) and finish treatment with D-R scores indicative of the emergence of object constancy (D-R 6; Blatt, Auerbach, & Aryan, 1998; Vermote et al., 2011).

Strengths. The ORI can be a useful aid in assessing the prevalence of an introjective or anaclitic orientation (P Axis) and in assessing several M Axis capacities, such as those for differentiation and integration, self-esteem regulation, and relationships and intimacy.

The ORI is easy to administer. It collects information about how a person understands the significant figures in his or her life. At the same time, it permits the assessment of structural dimensions of object relations that are not immediately evident and that reflect underlying levels of personality organization. Some raters without a psychodynamic or psychoanalytic background have been successfully trained to reliability in two days, both in Belgium and Israel (Luyten, personal communication to J. Auerbach, November 14, 2014). However, Diamond (personal communication to J. Auerbach, May 3, 2015) believes that a master's degree, clinical experience with psychodynamic theory, and a two-day training are necessary for competence in coding D-R.

Limitations. ORI scales are strongly tied to psychoanalytic object relations theories, and these ties may limit the generalizability and understandability of the measures in question for clinicians not conversant with these theories. The historical evolution and the scoring complexity of the ORI do not favor busy practitioners. To score, the 30- to 60-minute interview must be transcribed verbatim and then rated by a reliable coder; reliable coders are rare in routine practice. Clinical understanding of the findings requires understanding of the value of differentiating between the affective content of a significant-figure description and its level of psychological organization.

For more information, write to John Auerbach (auerbachjohn231@yahoo.com).

References

- Bender, D. S., Morey, L. C., & Skodol, A. E. (2011). Toward a Model for Assessing Level of Personality Functioning in DSM-5, Part I: A Review of Theory and Methods. *Journal of Personality Assessment, 93*, 332-346. doi: 10.1080/00223891.2011.583808
- Blatt, S. J., Bers, S. A., & Schaffer, C. E. (1993). *The assessment of self*. Unpublished research manual, Yale University.
- Besser, A., & Blatt, S. J. (2007). Identity consolidation and internalizing and externalizing problem behaviors in early adolescence. *Psychoanalytic Psychology, 24*, 126-149.
- Blatt, S. J. (1974). Levels of object representation in anaclitic and introjective depression. *Psychoanalytic Study of the Child, 29*, 107-157.
- Blatt, S. J. (2008). *Polarities of experience: Relatedness and self-definition in personality development, psychopathology, and the therapeutic process*. Washington, DC: American Psychological Association.
- Blatt, S. J., & Auerbach, J. S. (2001). Mental representation, severe psychopathology, and the therapeutic process. *Journal of the American Psychoanalytic Association, 49*, 113-159.
- Blatt, S. J., Auerbach, J. S., & Aryan, M. (1998). Representational structures and the therapeutic process. In R. F. Bornstein & J. M. Masling (Eds.), *Empirical Studies of*

- Psychoanalytic Theories: Vol. 8. Empirical Investigations of the Therapeutic Hour* (pp. 63-107). Washington, DC: American Psychological Association.
- Blatt, S. J., & Blass, R. (1996). Relatedness and self definition: A dialectic model of personality development. In G. G. Noam & K. W. Fischer (Eds.), *Development and vulnerabilities in close relationships* (pp. 309-338). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Blatt, S. J., Wein, S. J., Chevron, E. S., & Quinlan, D. M. (1979). Parental representations and depression in normal young adults. *Journal of Abnormal Psychology, 88*, 388-397.
- Blatt, S. J., Chevron, E. S., Quinlan, D. M., Schaffer, C. E., & Wein, S. (1988). *The assessment of qualitative and structural dimensions of object representations* (Revised Edition). Unpublished research manual, Yale University.
- Blatt, S. J. & Shichman, S. (1983). Two primary configurations of psychopathology. *Psychoanalysis and Contemporary Thought, 6*, 187-254.
- Diamond, D., Blatt, S. J., Stayner, D., & Kaslow, N. (1991). *Self-other differentiation of object representations*. Unpublished research manual, Yale University.
- Diamond, D., Kaslow, N., Coonerty, S. & Blatt, S. J. (1990). Change in separation-individuation and intersubjectivity in long-term treatment. *Psychoanalytic Psychology, 7*, 363-397.
- Heck, S. A., & Pincus, A. L. (2001). Agency and communion in the structure of parental representations. *Journal of Personality Assessment, 76*, 180-184.
- Huprich, S. K., Auerbach, J. S., Porcerelli, J. H., & Bupp, L. L. (in press). Sidney Blatt's Object Relations Inventory: Contributions and future directions. *Journal of Personality Assessment*.
- Huprich, S. K., Pouliot, G. S., Nelson, S. M., Pouliot, S. K., Porcerelli, J. H., Cawood, C. D., & Albright, J. J. (in press). Factor structure of the Assessment of Qualitative and Structural Dimensions of Object Representations (AOR) scale. *Journal of Personality Assessment*.
- Lowyck, B., Luyten, P., Verhaest, Y., Vandeneede, B., & Vermote, R. (2013). Levels of personality functioning in patients with personality disorders. *Journal of Personality*

Disorders, 27, 320-336.

Priel, B. (2005). Representations in childhood: A dialogical perspective. In J. S. Auerbach, K. N. Levy, & C. E. Schaffer (Eds.). *Relatedness, self-definition, and mental representation: Essays in honor of Sidney J. Blatt* (pp. 43-57). London, England: Routledge.

Quinlan, D. M., Blatt, S. J., Chevron, E. S., & Wein, S. J. (1992). The analysis of descriptions of parents: Identification of a more differentiated factor structure. *Journal of Personality Assessment*, 59, 340-351.

Vermote, R., Lowyck, B., Luyten, P., Verhaest, Y., Vertommen, H., Vandeneede, B., Corveleyn, J., & Peuskens, J. (2011). Patterns of Inner Change and Their Relation with Patient Characteristics and Outcome in a Psychoanalytic Hospitalization-Based Treatment for Personality Disordered Patients. *Clinical Psychology and Psychotherapy*, 18, 303-313.

The Adult Attachment Interview (AAI)

The AAI classification system developed by Mary Main and colleagues (Main & Goldwyn, 1998; George, Kaplan, & Main, 1985) examines the structural and discourse characteristics of autobiographical narratives about attachment experiences and relationships in adults. It is a semistructured clinical interview designed to elicit thoughts, feelings, and memories about early attachment experiences in order to elucidate the individual's attachment representations (i.e. internalized object relations) and current state of mind with respect to early attachment relationships and experiences (George, Kaplan, & Main, 1996).

The interview consists of 20 questions asked in set order with standardized probes. Individuals are asked to give five words that reflect their childhood relationships with each parent and then to provide episodic memories that support their semantic descriptions. They are also asked about how their parents responded to them when they were in physical or emotional distress and about experiences of rejection. In addition, there are questions about loss and physical and sexual abuse or other overwhelming or threatening

experiences. Other questions encourage individuals to reflect on how their childhood experiences with their parents have affected their adult personalities. The technique has been described as having the effect of “surprising the unconscious” (George et al., 1985) and allowing numerous opportunities for the interviewee to elaborate upon, contradict, or fail to support previous statements.

The AAI is transcribed verbatim, and trained coders score the transcripts to assess the nature of early experiences with attachment figures and current state of mind regarding attachment. Coders use subscales that assess (a) the rater’s inferences about the individual's experiences of parents in childhood (e.g., the extent to which parents were loving, rejecting, neglecting, involving, or role reversing); (b) the individual’s style of discourse and overall state of mind with respect to attachment (e.g. idealization, anger, derogation, passivity of thought, insistence on lack of memory, metacognition, coherence of transcript, and coherence of mind);; and (c) the individual's overall state of mind with respect to attachment (e.g., degree of derogation of attachment). Subscales are generally rated from 1 (absent or very low) to 9 (high) for the dimension in question. Several additional scales rate the extent to which subjects are unresolved with regard to loss or trauma (Main & Goldwyn, 1998).

The patterns of subscale ratings are then used to assign individuals to one of five primary attachment classifications:

1. Secure-autonomous, characterized by ready access to attachment-related memories that are expressed in a coherent, well-organized, fresh and spontaneous manner and with internally consistent and integrated portrayal of attachment relationships.
2. Dismissing, characterized by devaluing or idealizing states of mind with respect to attachment, with little corroborating evidence or with loss of recall of attachment-related memories and experiences.
3. Preoccupied, characterized by enmeshed or entangled states of mind with respect to attachment figures or by oscillation between positive and negative valuations of attachment figures.
4. Unresolved, characterized by lapses in the monitoring of reasoning and discourse,

such as highly implausible statements regarding the causes and consequences of traumatic attachment-related events, loss of memory for attachment-related traumas, or lapses into confusion and silence around discussion of trauma or loss, in response to questions about loss and abuse.

5. Cannot Classify, when there is oscillation between two or more opposing attachment states of mind (such as dismissing and preoccupied) or shifts in attachment strategy midway through the interview or in regard to different attachment figures (Main & Goldwyn, 1998).

The first three categories of adult attachment (secure, dismissing and preoccupied) correspond to the attachment patterns first identified in one-year-old children by Ainsworth (Ainsworth et al, 1978) and are considered to be organized categories because they involve specific identifiable and consistent strategies for regulating emotion in the context of attachment relationships. By contrast, the Unresolved and Cannot Classify categories, which were identified later by Main and her colleagues (Hesse, 1996; Hesse, 2010; Main & Weston, 1981) as a result of failure to classify a significant proportion of interviews in the organized categories, bespeak breaks and discontinuities in both behavioral and discourse strategies and failures of self and affect.

Main (Main, Hesse, & Goldwyn, 2008) has noted that the AAI interview and classification system parallels the process of analytic listening in that it involves paying attention to what individuals say, how they say it, and what they omit. The AAI classification system also provides linguistic markers to track defensive processes and modes of expressing and regulating affect that are useful to clinicians even if they do not have formal AAI training.

Strengths. The AAI can be a useful aid in the assessment of several relevant capacities of the M Axis (such as the relationship and intimacy capacity and the psychological self-observing and metacognitive capacity). Previous research has shown remarkable stability and predictive validity of the AAI (Water & Hamilton, 2000; Hesse, 2010). Parents' mental representation of attachment both before and after the birth of their child

has been found to predict later infant attachment status in over 18 international studies in a meta-analysis (van IJzendoorn, 1995). Thus, the secure and insecure patterns of attachment revealed on the AAI appear to provide a window on the nature of early parent-child relationships and tap into behavioral manifestations of normal and disturbed internalization of object relations (internal working models) that may substantially affect therapeutic discourse and process. The AAI status of patients has been found to be associated with therapeutic alliance, therapeutic process, and outcome (e.g., Bernecker, Ley, & Allison, 2014; Diamond et al., 1991, Diamond et al., 2003; Dozier, 1990; Dozier, Cue, & Bennett, 2001; Eagle, 2003; Eames & Roth, 2000; Rubion, Barker, Roth, & Fearon, 2000; Slade, 2008; Tyrell, Dozier, Teague, & Fallot, 1999), as well as with clinical and personality disorders. For example, borderline personality disorder has been associated in a number of studies primarily with the preoccupied and unresolved (disorganized) attachment categories (Bakersman-Kranenburg & van IJzendoorn, 2009; Diamond et al., 2003; Fonagy et al., 1996; Levy et al., 2006). In addition, other Axis I and Axis II disorders have been linked to insecure-disorganized attachment classifications, a pattern of finding that suggests that insecure and in particular disorganized attachments may be common developmental risk factors (see van IJzendoorn & Kranenburg, 2008, for a review). Finally, recent research has shown a movement from insecure to secure and disorganized to organized attachment status in borderline patients over the course of psychoanalytically oriented psychotherapy (e.g., Buchheim et al, 2014; Diamond et al., 2003; Levy et al., 1996; Levy et al., in press). Hence the AAI is a useful instrument to assess change in attachment representations over the course of psychodynamic treatment.

Limitations. Formal training to become an AAI coder is a lengthy process. The interviews must be audio-recorded and carefully transcribed according to a set of rules that capture both verbal utterances and breaks in the flow. AAI training requires a two-week-long workshop with a certified AAI trainer that involves learning to administer the interview and to understand the fundamentals of the coding system and its research and theoretical underpinnings, followed by independent coding of several sets of AAI

transcripts in order to achieve reliability. There are also limitations to the coding system in that it provides one overall attachment classification and does not take into account differences in the attachment representations of different figures (e.g., mother and father). Finally, the AAI classification system does not provide a dimensional measure of security or insecurity. However, the subscale ratings, particularly the coherence rating, which has been found in previous research to be the best predictor of attachment security (Waters, Treboux, Fyffe, & Crowell, 2001), are useful ratings to assess attachment dimensionally and to track incremental changes towards security in the course of treatment (Levy et al., 2006).

For more information, write to Diana Diamond (ddiamonda@gmail.com).

References

- Ainsworth, M. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the Strange Situation*. Hillsdale, NJ: Erlbaum.
- Bernecker, S.L., Levy, H. K., Ellison, W.D.(2014).Ameta-analysis of the relation between adult attachment style and the working alliance. *Psychotherapy Research*, 24, 1, 12-24.
- Bowlby, J. (1977). The making and breaking of affectional bonds: I. Aetiology and psychopathology in the light of attachment theory. *British Journal of Psychiatry*, 130, 201–210.
- Bakermans-Kranenburg, M. J., & van IJzendoorn, M. (2009). The first 10,000 Adult Attachment Interviews: Distributions of adult attachment representations in clinical and non-clinical groups. *Attachment & Human Development*, 11, 223–263.
doi:10.1080/14616730902814762
- Diamond, D., Clarkin, J. F., Levine, H., Levy, K., Foelsch, P., & Yeomans, F. (1999). Borderline conditions and attachment: A preliminary report. *Psychoanalytic Inquiry*, 19, 831–884.
- Diamond, D., Dozier, M. (1990). Attachment organization and treatment use for adults with serious psychopathological disorders. *Development and Psychopathology*, 2, 47–60.

- Diamond, D., Stovall-McClough, C., Clarkin, J., & Levy, K. N. (2003). Patient-therapist attachment in the treatment of borderline personality disorder. *Bulletin of the Menninger Clinic*, 67: 227-259.
- Dozier, M., Cue, K. L., & Barnett, L. (1994). Clinicians as caregivers: The role of attachment organization in treatment. *Journal of Consulting and Clinical Psychology*, 62, 793–800.
- Eagle, M. (2003). Clinical implications of attachment theory. *Psychoanalytic Inquiry*, 23, 12–27.
- Eames, V. and Roth, A. (2000). Patient attachment orientation and the early working alliance: A study of patient and therapist reports of alliance quality and ruptures. *Psychotherapy Research*, 10: 421-434.
- Fonagy, P., Leigh, T., Steele, M., Steele, H., Kennedy, R., Mattoon, G., et al. (1996). The relation of attachment status, psychiatric classification, and response to psychotherapy. *Journal of Consulting & Clinical Psychology*, 64, 22–31.
- Fonagy, P., Steele, M., Steele, H., & Target, M. (1998). Reflexive-function manual: Version 5.0 for application to the adult attachment interview. Unpublished manual, University of London.
- Hesse, E. (1996). Discourse, memory, and the Adult Attachment Interview: A note with emphasis on the emerging cannot classify category. *Infant Mental Health Journal*, 17, 4–11.
- Hesse, E. (2010). The Adult Attachment Interview: Protocol, method of analysis, and empirical studies. In J. Cassidy & P. Shaver (Eds.), *Handbook of attachment: Theory, research and clinical applications* (pp. 552–598). New York, NY: The Guilford Press.
- Levy, K. N., Meehan, K. B., Kelly, K. M., Reynoso, J. S., Weber, M., Clarkin, J. F., & Kernberg, O. F. (2006). Change in attachment patterns and reflective function in a randomized control trial of transference focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology*, 74, 1027–1040.
- Main, M., & Goldwyn, R. (1998). Adult attachment scoring and classifications system. Unpublished scoring manual. Department of Psychology, University of California,

Berkeley.

- Main, M., Goldwyn, R., and Hesse, E. (2003). *Adult Attachment Scoring and Classification System (Version 7.2)*. Berkeley, CA: University of California.
- Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. *Monographs of the Society for Research in Child Development*, 50, 66–104.
- Main, M., Hesse, E., & Goldwyn, R. Studying differences in language usage in recounting attachment history: An introduction to the AAI. In H. Steele & M. Steele, (Eds.), *Clinical Applications of the Adult Attachment Interview*. (pp. 31-68). New York: the Guilford Press.
- Main, M., & Weston, D. R. (1981). The quality of the toddler's relationship to mother and to father: Related to conflict behavior and the readiness to establish new relationships. *Child Development*, 52, 932–940.
- Slade, A. (2008). The implications of attachment theory and research for adult psychotherapy: research and clinical perspectives. In J. Cassidy and P. Shaver (Eds.), *Handbook of Attachment: Theory, Research and Clinical Applications*. (Second edition), pp. 762-783. New York: Guilford Press.
- Steele, H. and Steele, M. (2008). Ten clinical uses of the Adult Attachment Interview. In H. Steele and M. Steele (Eds.), *Clinical Applications of the Adult Attachment Interview*, pp. 3-30. New York: Guilford Press.
- Tyrrell, C. L., Dozier, M., Teague, G. B., & FalLOT, R. D. (1999). Effective treatment relationships for persons with serious psychiatric disorders: The importance of attachment states of mind. *Journal of Consulting and Clinical Psychology*, 67: 725-733.
- van IJzendoorn, M. H. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin*, 117, 387-403.
- van IJzendoorn, M. H., & Bakersman-Kranenburg, M. J. (2008). The distribution of Adult Attachment representations in clinical groups: A meta-analytic search for patterns of attachment in 105 AAI studies. In H. Steele & M. Steele (Eds.), *Clinical applications*

of the Adult Attachment Interview (pp. 399-426). New York, NY: Guilford Press.

Waters, E., and Hamilton, C. E. (2000). The stability of attachment security from infancy to adolescence and early adulthood: General introduction. *Child Development*, 71 (3): 678-683.

Waters, E., Treboux, D., Fyffe, C., & Crowell, J. (2001). Secure versus insecure and dismissing versus preoccupied attachment representation scored as continuous variables from AAI state of mind scales. New York: Stony Brook University, State University of New York.

The Adult Attachment Projective Picture System (AAP)

The AAP (George & West, 2001; George & West, 2012) is a free-response assessment of adult attachment status. The assessment is based on the analysis of "story" responses to a set of seven drawings depicting attachment scenes: four scenes of individuals alone and three of individuals in dyadic relationships. Individuals are instructed to tell a story for each stimulus, guided by a standard set of prompts that elicit story events, and characters' thoughts, and feelings. Administration is done on an individual basis in a private setting. Training is required for AAP administration and coding. Administrators do not need to be reliable AAP judges.

Coding examines the degree to which narratives reveal representational integration and attachment relationship attunement and sensitivity. Evaluations of responses to the alone stimuli include two dimensions: *agency of self* and *connectedness*. Agency of self evaluates capacity to view attachment relationships as fostering integrative or productive action. Connectedness evaluates the propensity toward making or sustaining relationships. Attunement and sensitivity are evaluated in the dyadic narratives on the basis of depictions of relationship *synchrony* that evidences the capacity for attachment-based contingent responsiveness or intimacy.

Defensive process coding provides important insight into differences among attachment classification groups. Every story is coded for three forms of defensive processes delineated in attachment theory (Bowlby, 1980; George & West, 2012): *deactivation*,

cognitive disconnection, and *segregated systems*. Deactivation shifts attention away from events, individuals, or affect (e.g., rejection). Cognitive disconnection fractures attention to the extent that relevant attachment information and affect become confusing or contradictory. Segregated systems blocks attention from painful distress, fear, rage, and helplessness.

Responses are also evaluated for *personal experience*, a dimension that demonstrates whether individuals are able to maintain self-other boundaries. The inclusion of personal experience material suggests absorption with distress or trauma.

The AAP is used to identify attachment group classifications and pathological mourning. There are four standard adult attachment patterns—three organized (secure or F, dismissing or Ds, preoccupied or E), and one dysregulated (unresolved or U). Secure attachment is characterized by substantial evidence of representational integration and value of attachment relationships. Dismissing and preoccupied attachment are characterized by defense patterns that interfere with integration but that help individuals engage in functional activities and relationships. Unresolved attachment is characterized by the failure to integrate or develop functional strategies in response to attachment trauma. Pathological mourning is evaluated from the presence of traumatic indicators in AAP responses and demonstrates risk for failure to complete the mourning process that is associated with heightened psychological and relationship distress (Bowlby, 1980).

The AAP has established construct validity and reliability. Concurrent validity was examined by comparing classification group designation using the AAP with the Adult Attachment Interview (AAI; George, Kaplan, & Main, 1984/1985/1996; Hesse, 2008), the international gold standard for assessing developmentally based adult attachment patterns. Validity and reliability studies in adults (ages 18+ years) are impressive. Studies show concordance rates in the two main construct validity studies ranged from 84-100 percent (Buchheim et al., 2006; Buchheim & George, 2011; George & West, 2001; George & West, 2012). Test-retest reliability is very good (George & West, 2012). The validity and interrater reliability of the AAP have been demonstrated by independent investigators (e.g., Webster & Hackett, 2007). The AAP has also been demonstrated as valid for use with

adolescents (≥ 13 years) (e.g., Aikins, Howes, & Hamilton, 2009).

Published case examples demonstrate how the AAP may prove useful clinically (George & Buchheim, 2014; Buchheim, Labek, Walter, & Viviani, 2013; Finn, 2011; Smith & George, 2012). The AAP has also been shown to be a useful resource for examining the neurophysiological attachment underpinnings of adult disorders (Buchheim et al., 2009; Buchheim, George, & Kächele, 2008).

The AAP can be useful for assessing several M and MA Axis functions, particularly those connected to defensive functioning and relational capacities.

Strengths. The AAP is a well-validated, user friendly, and economical measure that provides information about both the overarching classification patterns and underlying attachment processes needed to facilitate integrating developmental attachment theory into clinical work. The AAP can be used to evaluate course of treatment progress at different points and change over the course of time. Because the AAP is not a biographical task, it frees interviewees from taking ownership, so to speak, of attachment distress or trauma; by contrast, biographical assessment has been demonstrated problematic in attachment interview methodologies (Spieker, Nelson, DeKlyen, Jolley, & Mennet, 2011). The free-response methodology is straightforward, and administration is easy to learn. Individuals administering the AAP do not need to be trained in coding and classification. The AAP is easily used with non-English speaking interviewees, as has been demonstrated to date for European languages. Work examining use of the AAP in Asian and African languages is being explored.

Limitations. The assessment requires about 30-45 minutes and needs to be administered in an individualized (i.e., private) setting. The AAP stimuli are available through training so as to provide basic information about interpretation. Individuals wishing to code and classify their own transcripts must become reliable. The time it takes to achieve reliability is offset by making available reliable master judges who can code and classify cases.

For more information, see www.attachmentprojective.com.

References

- Aikins, J. W., Howes, C., & Hamilton, C. (2009). Attachment stability and the emergence of unresolved representations during adolescence. *Attachment and Human Development, 11*(5), 491-512. doi: 10.1080/14616730903017019
- Bowlby, J. (1980). *Attachment and loss: Vol. 3. Loss: Sadness and depression*. New York, NY: Basic Books.
- Buchheim, A., Erk, S., George, C., Kaechele, H., Ruchsow, M., Spitzer, M., . . . Walter, H. (2006). Measuring attachment representation in an fMRI environment: A pilot study. *Psychopathology, 39*, 144-152.
- Buchheim, A., & George, C. (2011). The representational, neurobiological and emotional foundation of attachment disorganization in borderline personality disorder and anxiety disorder. In J. Solomon & C. George (Eds.), *Disorganization of attachment and caregiving* (pp. 343-382). New York, NY: Guilford Press.
- Buchheim, A., George, C., Gündel, H., Heinrichs, M., Koops, E., O'Connor, M.-F., & Pokorny, D. (2009). Oxytocin enhances the experience of attachment security. *Psychoneuroendocrinology, 34*, 1417-1422.
- Buchheim, A., Labek, K., Walter, S., & Viviani, R. (2013). A clinical case study of a psychoanalytic psychotherapy monitored with functional neuroimaging. *Frontiers in Human Neuroscience, 7*.
- Finn, S. (2011). Use of the Adult Attachment Projective Picture System (AAP) in the middle of a long-term psychotherapy. *Journal of Personality Assessment, 93*, 427-433.
- George, C., & Buchheim, A. (2014). Use of the Adult Attachment Projective Picture System with a severely traumatized patient: A psychodynamic perspective. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.00865
- George, C., Kaplan, N., & Main, M. (1984/1985/1996). *The Adult Attachment Interview*.

University of California. Berkeley.

- George, C., & West, M. (2001). The development and preliminary validation of a new measure of adult attachment: The Adult Attachment Projective. *Attachment and Human Development, 3*, 30-61.
- George, C., & West, M. (2012). *The Adult Attachment Projective Picture System: Attachment theory and assessment in adults*. New York, NY: Guilford Press.
- Hesse, E. (2008). The Adult Attachment Interview: Protocol, methods of analysis, and empirical studies. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (2nd ed., pp. 552-598). New York, NY: Guilford Press.
- Smith, J. D., & George, C. (2012). Therapeutic assessment case study: Treatment of a woman diagnosed with metastatic cancer and attachment trauma. *Journal of Personality Assessment, 94*, 1-14.
- Spieker, S., Nelson, E. M., DeKlyen, M., Jolley, S. N., & Mennet, L. (2011). Continuity and change in unresolved classifications of Adult Attachment Interviews with low-income mothers In J. Solomon & C. George (Eds.), *Disorganized attachment and caregiving* (pp. 80-109). New York, NY: Guilford Press.
- Webster, L., & Hackett, R. K. (2007). A comparison of unresolved and resolved status and its relationship to behavior in maltreated adolescents. *School Psychology International, 28*, 365-378.

The Reflective Function Scale (RFS)

A group of clinician-researchers in the United Kingdom have developed this concept and instrument over the past twenty years (Fonagy, Target, Steele, & Steele, 1998). Emanating from an integration of frameworks from attachment theory, psychoanalysis, philosophy, and developmental affective neuroscience, *reflective function* or mentalization refers to the capacity to reflect on the self in terms of intentional mental states such as thoughts, feelings, and desires (Allen, Fonagy & Bateman, 2008; Fonagy et al., 2002). The capacity to mentalize is conceptualized as developing from secure attachment experiences

characterized by affective contingency and marked mirroring with caregivers; mentalization has critical implications for the maturation of adaptive relational capacities and affect regulation (Fonagy et al., 2002).

Mentalizing has been predominantly assessed with the Reflective Function (RF) scale, a performance-based measure, as applied to the Adult Attachment Interview (AAI) (Fonagy, Target, Steele, & Steele, 1998; see previous section) or the Parent Development Interview (Slade et al., 2004). An abbreviated version of the AAI has been recently developed to promote greater efficiency in assessing RF (Falkenstrom, Solbakken, Moller, Lech, Sandell, & Holmqvist, 2014).

The RFS is applied to transcripts of different sections of the AAI or PDI that specifically relate to interactions with attachment figures. Coders assess the extent to which the interviewee understands attachment-associated experiences in terms of mental states. The process of coding relies on distinguishing between “demand” and “permit” questions, with demand questions weighted more heavily. Demand questions directly probe for reflective functioning, whereas permit questions do not. Examples of demand questions include inquiries about which parent an interviewee felt closer to, experiences of parental rejection, interviewee interpretations of caregiver influence on his or her development, any life setbacks, interviewee reflections on the experience of death and loss, and interviewee reflections on the quality of the current relationship to parents and partner. Allocating ratings to the interviewee’s answer to each AAI question is dependent upon the coder’s identifying the presence and absence of qualitative markers of RF (the acknowledgment of opacity of mental states, separateness of minds between interviewer and interviewee, developmental aspects of mental states, and efforts to understand behavior in terms of mental states).

Answers to each question from the AAI are rated on an 11-point scale from 1 (systematic dismissal, derogation or hostility at any attempts at reflection; active avoidance of mentalization) to 9 (exceptional sophistication in the understanding of complex mental states; exceptional mentalization). A score of 5 is proposed as a normal level of mentalization and is given when convincing indications of a coherent model of the mind

are shown. A global score is obtained by individually weighting and totaling the ratings of the demand questions and permit questions, and allowing for a consideration of the interview as a whole.

A comprehensive review (Taubner et al., 2013) showed that the global RFS score can be assessed with good interrater reliability and has good stability across time. Further, all but one of the different demand questions from the AAI that are used to calculate the RFS score have incremental predictive validity of the global score. Interestingly, each of the five permit questions also contribute to the global score, even when controlling for the demand questions.

Studies with the RFS have shown that the measure is related to infant and adult attachment status. Furthermore, the RFS has been associated with a wide range of psychopathology, including borderline personality disorder, depression, eating disorders, panic disorder and psychosis. The RFS has also been related to trauma, deprivation in particular, and to posttraumatic stress disorder. Finally, studies also suggest that reflective functioning as measured with the RFS may moderate or mediate treatment outcome in various types of psychotherapy (for a review, see Fonagy & Luyten, in press; Katznelson, 2014).

Strengths. The RFS can be of aid for assessing both the reflective capacity of Axes M and MA and for assessing basic personality functions. Over the past two decades, perhaps the most generative new concept in psychoanalysis has been that of mentalizing. Studies with the RFS and other measures of reflective functioning (Luyten et al., 2012) promise to shed light on a key factor in personality development and the therapeutic process. The research base on the RFS is quite established and growing rapidly.

Limitations. The training program to reliably score the RFS is extensive and takes considerable time and practice. As a result, it has not been conducive to routine clinical application.

For more information, see www.annafreud.org.

References

- Allen, J., Fonagy, P., & Bateman, A. (2008). *Mentalizing in clinical practice*. Washington: American Psychiatric Press.
- Falkenstrom, F., Solbakken, O.A., Moller, C., Lech, B., Sandell, R., & Holmqvist, R. (2014). Reflective functioning, affect consciousness, and mindfulness: Are these different functions? *Psychoanalytic Psychology, 31*, 26-40.
- Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self*. New York, NY: Other Press.
- Fonagy, P., Target, M., Steele, M. & H. Steele, (1998). *Reflective Functioning Manual: Version 5. For application to Adult Attachment Interviews*. London: University College.
- Katznelson, H. (2014). Reflective functioning: A review. *Clinical Psychology Review, 34*(2), 107-117.
- Luyten, P., Fonagy, P., Lowyck, B., & Vermote, R. (2012). The assessment of mentalization. In A. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (pp. 43-65). Washington, DC: American Psychiatric Association.
- Slade, A., Aber, J. L., Bresgi, I., Berger, B., & Kaplan, C. A. (2004). The Parent Development Interview - Revised. *Unpublished protocol. The City University of New York, New York*.
- Taubner, S., Horz, S., Fischer-Kern, M., Doering, S., Buchheim, A., & Zimmermann, J. (2013). Internal structure of the Reflective Functioning Scale. *Psychological Assessment, 25*(1), 127-135.

The Quality of Object Relations (QOR)

Created by Piper and colleagues in the 1970s, this scale was modified by Høglend (1993). From a single assessment interview, quality of object relations are rated on a spectrum from primitive to mature. It is largely psychodynamic in its focus with emphasis on recurring interpersonal patterns in development and current life.

The QOR assessment is standardized and is based on a one- to two-hour interview in which the rater first asks for a history of the significant relationships of the client, reported in a spontaneous way, and then assesses the quality of object relations on an 8-point scale, with scale points described as follows (Høglend, QOR Manual, unpublished):

- 8 – 7: A history of most relationships characterized by stability, gratification and mutuality. Others are seen as whole, autonomous persons.
- 6 – 5: Recent interpersonal functioning may be poorer, but the patient can give detailed examples from at least one earlier important high quality relationship, in adulthood, adolescence, or childhood. Conflictual feelings may be seen towards same-sex persons and fears of loss may be present toward opposite-sex persons.
- 4 – 3: Mostly stable but less gratification and mutuality in most important relations. Passivity, dependency, or need to control others are predominant, or separation anxiety. Difficulty describing others as unique individuals.
- 2 – 1: Mostly unstable relations with little valued persons. Others are seen as need-gratifying objects. Stable, overly dependent relations only with parental objects.

Two additional scales of the QOR, with points defined as above, entail history of intimate sexual relationships and history of friendships.

Validation has been reported by Piper et al. (1999). A recent study showed the modifying impact of the quality of object relations on the results of treatment in regard to self-concept (Lindfors et al. 2013). This finding converges with the findings of Høglend et al. (2011) about differences in the impact of transference interpretations depending on level of object relations (see “Psychodynamic Functioning Scale” described below).

Strengths. The QOR can be a useful aid in the assessment of the relationships and intimacy capacity of the M axis, and can give useful information also for the assessment of the P Axis. As may be seen in the description of the scale points above, there are many

overlaps between the QOR, the SCORS, and the ORI (see above). These latter measures are more precisely developed; however, the QOR has the great advantage of being based on a single semi-structured clinical interview, and not requiring specific training for the clinician-rater beyond careful study of the manual.

Limitations. Emphasis on interpersonal and object relations functioning need often to be integrated with other aspects of psychodynamic understanding, such as defensive operations, identity, reflective functioning, values and so on. Moreover, other instruments are better developed and researched than this assessment tool. Studies of convergent validity with other measures of interpersonal relations would be valuable.

For more information, write to Per Høglend (hoglend@medisin.uio.no).

References

- Azim, H, Piper W, Segal P, Nixon G, Duncan S. (1991). The Quality of Object Relations Scale, *Menninger Clinic, Bulletin*, 55: 323-343.
- Høglend, P. Hersoug, A.G., Bøgwald, K.P., Amlo, S., Marble, A., Sørbye, Ø., Røssberg, J.I., Ulberg, R., Gabbard, G. & Crits-Christoph, P. (2011). Effects of Transference Work in the Context of Therapeutic Alliance and Quality of Object Relations. *Journal of Consulting and Clinical Psychology* 79, 697–706.
- Huprich, S. K., & Greenberg, R. P. (2003). Advances in the assessment of object relations in the 1990s. *Clinical Psychology Review*, 23, 665–698.
- Lindfors, O. , Knekt, P. and Virtala, E. (2013) Quality of object relations modifies the effectiveness of short- and long-term psychotherapy on self-concept. *Open Journal of Psychiatry*, 3, 345-350.
- Piper, W. E., McCallum, M., & Joyce, A. S. (1993). Manual for assessment of quality of object relations. Unpublished manuscript.
- Piper W.E., Duncan SC. (1999) Object relations theory and short-term dynamic psychotherapy: findings from the Quality of Object Relations Scale. *Clin Psychol Rev.* 1999, 19:669-85.

The Plan Formulation Method (PFM)

The PFM (Curtis, Silberschatz, Sampson, & Weiss, 1994; Curtis & Silberschatz, 2005, 2007) is a psychodynamic formulation method based on Sampson and Weiss' control-mastery theory (Weiss, Sampson, and the Mount Zion Psychotherapy Research Group, 1986; Weiss, 1993). There are five interrelated components in the plan formulation:

- the **patient's** adaptive conscious and unconscious short-term and **long-term goals**
- the **unconscious pathogenic beliefs** or schemas developed as a consequence of shock and strain trauma that impede goal attainment
- the **adverse or traumatic childhood experiences** that led to pathogenic beliefs or schemas
- the **patient's unconscious tests of the therapist** in order to disconfirm pathogenic beliefs and expectations
- **new information or insight** that might help the patient to overcome her or his pathogenic beliefs or schemas and their consequences.

Excellent reliability and validity data have been reported. Trained clinicians show a high level of interrater reliability with regard to dimensions of the model such as the patient's unconscious goals, pathogenic beliefs (Curtis & Silberschatz, 2005, 2007). Research on detailed study of audio recordings has also shown that plan formulations can be used to predict therapist responses and interventions that are helpful to patients and those that are not (e.g., Silberschatz, Fretter & Curtis, 1986; Silberschatz & Curtis, 1993; for review, see Silberschatz, 2005).

Strengths. The PFM may be a useful aid for the assessment of P Axis pathogenic beliefs and the capacity for relationships and intimacy of the M Axis. This method provides systematic measures for creating a dynamic formulation (Eells 1997, 2009). The major strength of the PFM is its high degree of reliability and its validity. The method has yielded strong reliability data for psychoanalytic cases, brief dynamic psychotherapy, cognitive therapy, and crisis intervention (Silberschatz, 2005). Research has also shown that the formulations derived from the PFM can be used to assess the suitability or

responsiveness of therapist interventions (Silberschatz, 2005).

Limitations. As with any complex approach to formulation, mastery of the PFM requires some investment of time and supervision. Most practitioners and many investigators may not have the resources to learn this method, and this problem has undoubtedly limited its widespread use for research or for training purposes.

For more information, write to George Silberschatz (George.Silberschatz@ucsf.edu).

References

- Curtis, J. T., Silberschatz, G., Sampson, H. & Weiss, J. The Plan Formulation Method. *Psychotherapy Research*, 1994, 4, 197-207
- Curtis, J. T. & Silberschatz, G. The assessment of pathogenic beliefs. In G. Silberschatz (Ed.), *Transformative relationships* (pp. 69-92). New York: Routledge, 2005
- Curtis, J. T. & Silberschatz, G. Plan formulation method. In T. D. Eells (Ed.), *Handbook of psychotherapy case formulation* (Second Edition, pp. 198 -220). New York: Guilford, 2007.
- Eells, T. D. (2009). Contemporary themes in case formulation. In Sturmey, Peter [Ed]. *Clinical case formulation: Varieties of approaches*: 293-315. Hoboken, NJ, John Wiley & Sons .
- Eells, T. D [Ed]. (1997). *Handbook of psychotherapy case formulation*. New York, Guilford Press.
- Silberschatz, G. An overview of research on control-mastery theory. In G. Silberschatz (Ed.), *Transformative relationships* (pp. 189-218). New York: Routledge, 2005
- Silberschatz, G. & Curtis, J. T. Measuring the therapist's impact on the patient's therapeutic progress. *Journal of Consulting and Clinical Psychology*, 1993, 61, 403-411
- Silberschatz, G., Fretter, P. B., & Curtis, J. T. How do interpretations influence the process of psychotherapy? *Journal of Consulting and Clinical Psychology*, 1986, 54, 646-652
- Weiss, J., Sampson, H., and the Mount Zion Psychotherapy Research Group, The

Psychoanalytic Process. New York: Guilford Press, 1986.

The Core Conflictual Relationship Theme (CCRT)

Luborsky and Crits-Christoph (1990) defined the CCRT as a method of understanding an individual's way of relating to others and how this manner of interaction plays out in the relationship with the psychotherapist. The method requires first identifying, in recorded text, relationship episodes (REs). Luborsky noted that patients' narratives showed a considerable redundancy in the expression of a core Wish (W), occasionally more than one, a typical anticipated or actual Response from Other persons (RO) in the narrative, and then a subsequent Response from the Self (RS) to the reaction of the other.

Initially, Luborsky identified these three components for a given relationship episode by clinical inspection. The specifications of W, RO, and RS were idiographic—that is, described by raters in words they thought most appropriate, without reference to any established categories. When multiple judges were used, there were difficulties in assessing the degree of agreement between them. To solve this, agreement judges were used. With these procedures, it was possible to establish consensus formulations. Such formulations suffered, however, from both the complexity of the procedures and the difficulty in estimating the reliability of the results. Therefore, Luborsky's research group developed standard categories (about 30) for each of the three components (W, RO, and RS). Later, a German-speaking research group developed a different categorization of the CCRT component, the Core Conflictual Relational Themes-Leipzig/Ulm (CCRT-LU; Albani et al., 2002)

Typically, two or three early sessions from early in a treatment are recorded, and REs are identified until 10 have been obtained. Each RE is subdivided into thought units, which are essentially sentences, and each sentence is classified as a W, an RO or an RS. The most frequent W, RO, and RS (if any) for each RE is designated and counted for each of the ten REs. The CCRT for the patient is defined as simply the sum of the most frequent W, RO, and RS, across the 10 relationship episodes (REs).

Luborsky and Crits-Christoph (1990) classify RO and RS as to whether the response is

positive or negative. This is relevant because psychological conflict is initiated by wishes that threaten to provoke negative responses from others or from the self. Furthermore, the degree to which people are able to arrange for positive responses to their wishes directly reflects their adequacy in handling life. In studying the Penn Psychotherapy sample of 33 cases, the authors found that early in treatment negative attitudes of others and of self (RO and RS) tended to be much more frequent than positive ones. In the course of good-outcome treatment, judged by two independent measures of improvement, the negative attitudes substantially decreased, and the positive ones substantially increased, but there was no significant change in wishes.

Strengths. The CCRT may be a useful aid for the assessment of pathogenic beliefs about oneself and other people (P axis) and of the capacity for relationships and intimacy on the M Axis. The positive and negative assessment provides a method of systematically assessing what patients experience and what kinds of changes occur over the course of treatment.

The accuracy of interpretations has also been assessed on the basis of how much the Ws, ROs, and RSs are characterized in the interventions. Changes in the patients' self-awareness may also be reflected in their awareness of the CCRT components. These indices constitute a rough grid that can be applied to therapy recordings and thereby permit systematic assessment.

Limitations. First, the CCRT requires considerable time and training to score. Second, Luborksy and collaborators conceived of the CCRT in terms of a sequenced pattern of reactions by the patient (or other person). Yet, the CCRT as operationally defined does not necessarily reflect the patterns of the patient's concerns (see example in Waldron, 1995, p. 400).

For more information, write to Paul Crits-Cristoph (crits@mail.med.upenn.edu).

References

- Albani, C., Benninghofen, D., Blaser, G., Cierpka, M., Dahlbender, R. W., Geyer, M. & Kächele, H. (1999). On the connection between affective evaluation of recollected relationship experiences and the severity of psychic impairment. *Psychotherapy Research*, 9(4), 452-467.
- Albani, C., Pokorny, D., Blaser, G., Gruninger, S., König, S., Marschke, F., et al. (2002). Reformulation of the core conflictual relationship theme (CCRT) categories: The CCRT-LU category system. *Psychotherapy Research*, 12, 319-338.
- Drapeau, Martin; Perry, J. Christopher. (2009). The Core Conflictual Relationship Themes (CCRT) in borderline personality disorder. *Journal of Personality Disorders*. 23: 425-431.
- Hörz, S., Clarkin, J. F., Stern, B. L., & Caligor, E. (2012). The Structured Interview of Personality Organisation (STIPO): An instrument to assess severity and change of personality pathology. In R. Levy, S. Ablon & H. Kächele (Eds.), *Psychodynamic Psychotherapy Research* (pp. 571-592). New York: Humana.
- Luborsky, L., Crits-Christoph, P., Mintz, J. & Auerbach, A. (1988). *Who Will Benefit from Psychotherapy? Predicting Therapeutic Outcomes*. New York: Basic Books.
- Luborsky L & Crits-Christoph P (1990). *Understanding Transference: the CCRT Method*. New York: Basic Books.
- Waldron, S. (1995) Book Review: *Understanding Transference: The CCRT Method*. by Lester Luborsky and Paul Crits-Christoph. New York, Basic Books, 1990, *Psychoanalytic Quarterly*, 64:398-402.

The Defense Mechanisms Rating Scale (DMRS)

J. C. Perry, with his colleagues, came to the description of defense mechanisms as a product of his efforts to develop systematic methods of case formulation and to determine the differences between the functioning of borderline patients and others (Perry & Høglend, 1998). By classifying defenses according to Eriksonian stages of development, Perry was able to develop an index of overall defense functioning.

The DMRS manual (Perry, 1990; Perry & Kardos, 1995; Perry, 2001; Lingardi &

Madeddu, 2002) describes how to identify 30 individual defense mechanisms in videotaped or audiotaped sessions or transcripts. The manual presents a definition of each defense, a description of how the defense functions, a guide for the differential diagnosis of each defense and a 3-point scale. Each scale is clearly identified, with specific examples of (0) no use, (1) probable use and (2) definite use of the defense. The examples provide prototypical instances of the defense that expand and complement the formal definitions.

In the DMRS there are 7 defense levels arranged hierarchically, with each defense assigned to a particular level. The defense levels are characterized as follows in descending order of health:

7. High Adaptive Level (also called “Mature”): affiliation, altruism, anticipation, humor, self-assertion, self-observation, sublimation, suppression
6. Obsessional: isolation, intellectualization, undoing
5. Other Neurotic: repression, dissociation, reaction formation, displacement
4. Minor Image-distorting (also called “Narcissistic”): omnipotence, idealization (of self and of other people), devaluation
3. Disavowal: denial, projection, rationalization, autistic fantasy
2. Major-Image-distorting (also called “Borderline”): splitting of others’ image, splitting of self-image, projective identification
1. Action: acting-out, passive aggression, help-rejecting complaint

This ordering is based on a series of empirical studies (reviewed in Perry, 1993). Not included in the manual, but included in an appendix, are the psychotic defenses.

The rater identifies each use of the defense as it occurs, bracketing the part of the text in which it operates. After the completion of the ratings, the number of times each defense was identified in the text is divided by the total instances of all defenses to produce a percentage score for each defense. Then the total percentage of defenses at each level forms the basis for a “defense profile” that represents the nature of the patient’s functioning and that can be compared with earlier or later functioning in the course of treatment.

All of the defense scores are summarized by an Overall Defensive Functioning (ODF)

score (Perry & Høglend, 1998), an average of the scores for all scored defenses. In clinical samples based on whole interviews, scores usually range between 2.5 and 6.5.

Approximate reference scores for ODF are as follows:

1. Scores below 5.0 are associated with personality disorders, severe depression, or borderline conditions;
2. Scores between 5.0 and about 5.5 are associated with neurotic character and symptom disorders;
3. Scores from 5.5 to 6.0 are associated with average healthy-neurotic functioning; and
4. Scores above 6.0 are associated with superior functioning.

A Q-sort version of the instrument has been recently developed and validated (Di Giuseppe, Perry, Petraglia, Janzen, & Lingiardi, 2014).

The DMRS served as the basis for the DSM-IV Defensive Functioning Scale (DFS; American Psychiatric Association, 1994). Reliability and validity of the DFS has been reported in several studies (e.g., Blais, Conboy, Wilcox, & Norman, 1996; DeFife & Hilsenroth, 2005; Perry et al, 1998; Porcerelli, Cogan, Kamoo, & Miller, 2010; Porcerelli, Cogan, Markova, Miller, & Mickens, 2011). It has been implemented in large studies on personality, psychopathology, and defense mechanisms, as well as on changes in defense mechanisms during psychotherapy.

Strengths. The DMRS serves well to assess defensive functioning according to the M and MA Axis and in the context of the assessment of the level of personality organization (P Axis).

The psychometric reliability of the DMRS and its capacity to discriminate different disorders and levels of functioning make it an excellent instrument for the assessment of patients in psychodynamic psychotherapy (Perry et al., 2009). It also provides a general measure of health-sickness and can reflect changes in this dimension during the course of therapy. Several such studies have shown interesting results supporting the beneficial effects of the work on defenses in psychotherapy (Perry & Bond, 2012).

Limitations. The DMRS requires several days of specific training to achieve reliability, and the assessment is time consuming (approximately 2-3 hours for a 50-minute session). Moreover, there is no complete agreement on the definition and the categorization of some defenses (for example, dissociation). The amount of time needed for the assessment can be substantially reduced with the Q-sort version.

The DFS gives a less detailed, precise and reliable picture of patients' functioning than does the DMSR.

For more information, write to John Cristopher Perry (jcristopher.perry@mcgill.ca).

References

- Berney, S., de Roten, Y., Beretta, V., Kramer, U., & Despland, J.N. (2014). Identifying Psychotic Defenses in a Clinical Interview. *Journal of Clinical Psychology, 70*, pp. 428-439.
- Blais, M. A., Conboy, C. A., Wilcox, N., & Norman, D. K. (1996). An empirical study of the DSM-IV Defensive Functioning Scale in personality disordered patients. *Comprehensive Psychiatry, 37*, 435-440.
- Bond, M., & Perry, C.J. (2004). Long-term changes in defense styles with psychodynamic psychotherapy for depressive, anxiety, and personality disorders. *American Journal of Psychiatry, 161*, pp. 1665-1671.
- DeFife, J. A., & Hilsenroth, M. J. (2005). Clinical utility of the Defensive Functioning Scale in the assessment of depression. *Journal of Nervous and Mental Disease, 193*, 176-182.
- Di Giuseppe, M.G., Perry, J.C., Petraglia, J., Janzen, J., & Lingardi, V. (2014). Development of a Q-sort version of the Defense Mechanism Rating Scales (DMRS-Q) for clinical use. *Journal of Clinical Psychology, 70*, pp. 452-465.
- Lingardi, V., Lonati, C., Delucchi, F., Fossati, A., Vanzulli, L., & Maffei, C. (1999). Defense mechanisms and personality disorders. *Journal of Nervous and Mental*

- Disease, 187, 224-228.
- Perry, J.C. (1990). *Defense Mechanism Rating Scales (DMRS)* (5th edition). Cambridge, MA.
- Perry, J.C. (2014). Anomalies and specific functions in the clinical identification of defense mechanisms. In *Journal of Clinical Psychology*, 70, pp. 406-418.
- Perry, J.C. & Bond, M. (2012). Change in defense mechanisms during long-term dynamic psychotherapy and five-year outcome” *American Journal of Psychiatry*, 169, 916-925.
- Perry, J. C., Høglend, P., Shear, K., Vaillant, G. E., Horowitz, M., Kardos, M. E., & Kagen, D. (1998). Field trial of a diagnostic Axis for defense mechanisms for DSM-IV. *Journal of Personality Disorders*, 12, 56-68.
- Perry, J. Christopher, & Høglend, Per. (1998). Convergent and discriminant validity of overall defensive functioning. *Journal of Nervous and Mental Disease*. 186, 529-535.
- Porcerelli, J. H., Cogan, R., Kamoo, R., & Miller, K. (2010). Convergent validity of the Defense Mechanisms Manual and the Defensive Functioning Scale. *Journal of Personality Assessment*, 92(5), 432-438.
- Porcerelli, J. H., Cogan, R., Markova, T., Miller, K., & Mickens, L. (2011). The Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition Defensive Functioning Scale: A validity study. *Comprehensive Psychiatry*, 52, 225-230.

The Defense Mechanisms Manual (DMM)

The DMM was developed to assess three developmentally anchored defenses – denial, projection, and identification – applied to TAT stories (Cramer, 1991). The defenses are represented by seven different facets that are coded each time they occur within a story. Total scores or relative scores (the percentage of a defense in relation to the total number of defenses) can be used clinically and for research purposes.

An interrater reliability of .80 has been reported for the DMM (Meyer, 2004). The DMM has shown adequate stability over a 3-year period (Cramer, 1998). The validity of

the DMM has been supported through cross-sectional and longitudinal studies of child, adolescent, and young adult development, with these studies showing differentiation of diagnostic groups, of levels of personality organization, and of change following intensive psychodynamic treatment. The impact of stress on defense use has also been studied (Cramer, 2006; Porcerelli et al., 2010).

Strengths. The DMM can be a useful aid for the assessment of defensive functioning of the MC Axis, when Thematic Apperception Test materials are available. The DMM is a reliable measure and its validity is supported by cross-sectional, longitudinal, and experimental studies.

Limitations. The DMM includes only three defense mechanisms and the defense of identification may be less useful than are other defenses in the assessment of adults because its use diminishes after late adolescence. Administration of the TAT, creation of a transcript from the TAT stories, and use of the DMM rating system would take several hours in routine practice.

For more information, write to John Porcerelli (jporcer@med.wayne.edu).

References

- Cramer, P. (1991). *The development of defense mechanisms: Theory, research, and assessment*. New York, NY: Springer-Verlag.
- Cramer, P. (1998). Freshman to senior year: A follow up study of identity, narcissism and defense mechanisms. *Journal of Research in Personality*, 32, 156-172.
- Cramer, P. (2006). *Protecting the self: Defense mechanisms in action*. New York, NY: Guilford.
- Meyer, G. (2004). The reliability of the Rorschach and TAT compared to other psychological and medical procedures: An analysis of systematically gathered evidence. In M. Hilsenroth and D. Segal (Eds.) *Personality assessment: Volume 2*.

Comprehensive handbook of psychological assessment (pp. 315-342). Hoboken, NJ: Wiley.

Porcerelli, J. H., Cogan, R., Kamoo, R., & Miller, K. (2010). Convergent validity of the Defense Mechanisms Manual and the Defensive Functioning Scale. *Journal of Personality Assessment*, 92(5), 432-438.

Part Two: Self-reports

The Minnesota Multiphasic Personality Inventory-2nd Edition (MMPI-2)

Hathaway and McKinley (1940) developed the Minnesota Multiphasic Personality Inventory (MMPI) as an empirically based self-report instrument from a pool of statements found in patient records and other sources. Their goal was to develop a simple, objective test of psychopathology based on criterion groups. To this day, it is the only self-report measure of psychopathology based on actual cases. In 1989, the MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) was released, and it was based on the same empirical methodology described below for the construction of the original MMPI but used a more representative normative sample than the 1930s Minnesota medical outpatients who constituted most of the non-clinical participants in the initial control samples. There is also an adolescent version of the test, the MMPI-A, that was normed for persons between the ages of 14 and 18 (Butcher, Williams, Graham, Archer, Tellegen, Ben-Porath, & Kaemmer, 1992). More recently, the MMPI-2-Restructured Form (RF) (Tellegen, Ben-Porath, McNulty, Arbisi, Graham, & Kaemmer, 2003) was constructed on the basis of the MMPI-2 to minimize the statistical and clinical heterogeneity of the classical MMPI clinical scales. Although interpretation of the MMPI-2-RF is therefore more straightforward than is interpretation of the MMPI or MMPI-2, it is commonly thought, as discussed in this section, that the homogeneity of the MMPI-2-RF's clinical scales also decreases the psychodynamic content that was tapped by the psychometrically and clinically messier scales of the original measure.

Hathaway and McKinley (1940) developed scales on the basis of which items reliably differentiated clinical or criterion groups from a nonclinical group. For example, those items that differentiated people diagnosed with hypochondriasis from both medical patients and nonclinical individuals became the Hypochondriasis (*Hs*)

scale (1). In a similar manner, they went on to develop the Depression (*D*) scale (2), the Hysteria (*Hy*) scale (3), the Psychopathic Deviate (*Pd*) scale (4), the Masculinity-Femininity (*Mf*) scale (5), the Paranoia (*Pa*) scale (6), the Psychasthenia (*Pt*) scale (7), the Schizophrenia (*Sc*) scale (8), and the Hypomania (*Ma*) scale (9). Because the MMPI clinical scales are not pure measures of the constructs for which they are named, it is preferable to designate them by the number assigned to them (e.g., 1, rather than Hypochondriasis or *Hs*), but the scale names or abbreviations are often used because of the descriptive information they contain. Hathaway and McKinley also developed scales to measure validity and bias, specifically, the Lie (*L*) scale, which measures unsophisticated and rigid attempts to present oneself in a positive manner, the Infrequency (*F*) scale, measuring rare psychopathology and symptom exaggeration, and the Correction or Defensiveness (*K*) scale (McKinley, Hathaway & Meehl, 1948), which measures more subtle indications of defensiveness and social desirability and which is added, in whole or in part, to particular scales on which symptoms are likely to be underreported (i.e., scales 1, 4, 7, 8, and 9). Later, Drake (1946) developed a Social Introversion (*Si*) scale (0). The MMPI-2 yields an interpretable profile because the scores of its nonclinical samples are converted to T scores with a mean of 50 and a standard deviation of 10. On most MMPI-2 scales, T score of 65, 1.5 standard deviations above the nonclinical mean, is considered a clinically significant elevation.

Although the MMPI-2 was not developed as a psychodynamically oriented test, it can be useful in helping to determine PDM-2 diagnoses. The *L* scale can be an indication of primitive defenses such as denial and splitting (Gordon, Stoffey, & Bottinelli, 2008). The MMPI-2 basic clinical scales can measure traits to help in the assessment of the PDM-2 Personality Disorders: Scale 1 (*Hs*) relates to the Somatizing personality, scale 2 (*D*) relates to the Depressive personality, scale 3 (*Hy*) relates to the Hysterical (Histrionic) personality, scale 4 (*Pd*) relates to the Psychopathic and Narcissistic personalities, scale 6 (*Pa*) relates to the Paranoid personality, scale 7 (*Pt*) relates to the Anxious, Phobic or Obsessive-compulsive

personality, scale 8 (*Sc*) relates to the Schizoid personality, and scale 9 (*Ma*) relates to the Narcissistic, Psychopathic, and Hypomanic personalities. Sadistic personality and Sadomasochistic personality often have elevations on scales 4 (*Pd*), 8 (*Sc*), and 9 (*Ma*), as well as on the Emotional Alienation (*Sc2*) subscale of scale 8 (*Sc*). Masochistic personality often has elevations on scales 3 (*Hy*) and 7 (*Pt*), as well as on the Naivete (*Pa3*) subscale of scale 6 (*Pa*). Dependent personality often has elevations in scales 2 (*D*) and 7 (*Pt*). Counterphobic personality often has elevations in scales 4 (*Pd*) and 9 (*Ma*). Finally, Dissociative personality often has elevations on scales 3 (*Hy*) and 8 (*Sc*) (Gordon, 2007). Although individual scales are sometimes evaluated, they have usually been interpreted, in part because of the heterogeneity of scales but for other reasons as well, on the basis of two- and three-point code configurations. For example, elevations on scales 1 (*Hs*) and 3 (*Hy*) is considered a 1-3/3-1 profile, indicative of a high degree of somatic concern and, in conjunction with a low score on scale 2 (*D*), is often found in conversion disorder or classical hysteria. Two- and three-point codes are the best researched and mostly frequently used configurational components of MMPI interpretation.

The MMPI-2 can also be used to help determine the level of personality organization. The Goldberg (1965) formula has been shown to be useful in differentiating neurotic, borderline and psychotic levels of personality organization from each other (Gordon, Stoffey, & Bottinelli, 2008). The Goldberg formula is calculated, using T scores, as $(L + Pa + Sc) - (Hy + Pt)$, with values of greater than 65 indicating borderline or psychotic personality organization and values of less than 35 indicating neurotic personality functioning.

Strengths: To date, more than 19,000 books and articles have been published on the MMPI instruments (Butcher & Williams, 2009), and the test is used in many settings around the world. In fact, the MMPI-2 is the most used self-report measure of psychopathology (Camara, Norton, & Puente, 2000). The MMPI-2 is about 80-90% accurate in assessing personality traits and defensiveness, (p.102, Butcher, et al. MMPI-2 Manual, 1989), and Bram and Peebles (2014) still consider the MMPI-2,

TAT, Rorschach and WAIS as the core battery of personality assessment. The MMPI-2 has grown to have more than 15 scales and formulas for determining response bias and over 150 scales of psychopathology. Although the MMPI-2 is long, with 567 items, the “short” form, which is the first 370 items, produces all the main validity and clinical scales and subscales. Anyone with at least a sixth grade reading ability can take the MMPI-2.

The MMPI-2 measures rather stable psychopathology. Spiro, et al. (2000), using the MMPI-2 with 1,050 normal men over a period of 5 years between administrations, found a median stability index of .68, and a range of .56 to .86. The MMPI-A has similar reliability, .65 - .84 for the clinical scales, after a one-week interval (Butcher, et al. 1992).

Because the MMPI-2 measures stable personality traits, insensitive to low-dose treatments, it is rarely used as a dependent measure in treatment outcome research. However, Gordon (2001) found that, after an average of three years of psychoanalytic treatment, most of the clinical scales were significantly reduced, and the Ego Strength (*Es*) scale (Barron, 1953) was significantly increased.

Limitations: The MMPI-2 is limited by its methodology, i.e. depending on an individual’s response to “True” or “False” written questions. The MMPI-2 does not probe further on a hint of a symptom or a defense or use branching logic to rule out certain diagnoses on obtaining certain information. For example, it does not ask questions such as "How long have you been depressed?" or "Did some bad things happen to you before you became depressed?" or "Do you have periods of feeling very energetic after feeling depressed?" Such follow-up questions are often valuable in differentiating between depression from Dysthymia, or Posttraumatic Stress Disorder, or a Bipolar Disorder. No matter how much an item pool is improved, a pencil-and-paper test can have just so much validity. Also, at times it makes sense for a client to fake to look good (e.g. job screening, custody evaluations, etc.), and in such instances the MMPI-2 (as with all self-reports) may only show a false negative,

that is a seemingly normal-looking profile. However, the MMPI-2 validity scales are excellent and it is the best self-report instrument available for detecting faking to look good or faking to look bad (Hopwood & Bornstein, 2014).

Moreover, the MMPI is a long test that requires about 45-60 minutes to complete the 370 item form and about 60-90 minutes to complete the 567 item test, and it must be administered and interpreted by a psychologist with appropriate training. The more recent form, the MMPI-2-RF has high item internal consistency, but may lack clinical sensitivity (Gordon, Stoffey, & Perkins (2013).

For more information, write to Robert M. Gordon (rmgordonphd@gmail.com)

References

- Barron, F. (1953). An ego-strength scale which predicts response to psychotherapy. *Journal of Consulting Psychology, 17*, 327-333.
- Bram, A. D., & Peebles, M. J. (2014). Psychological testing that matters: Creating a road map for effective treatment. American Psychological Association, Washington, D.C.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). MMPI-2: Manual for administration and scoring. University of Minnesota Press, Minneapolis, MN.
- Butcher, J. N. and Williams, C. L. (2009). Personality Assessment with the MMPI-2: Historical Roots, International Adaptations, and Current Challenges. *Applied Psychology: Health And Well-Being, 1* (1), 105–135
doi:10.1111/j.1758-0854.2008.01007.
- Butcher, J.N., Williams, C.L., Graham, J.R., Tellegen, A., Ben-Porath, Y.S., Archer, R.P. et al. (1992). Manual for administration, scoring, and interpretation of the Minnesota Multiphasic Personality Inventory for Adolescents: MMPI-A. Minneapolis, MN: University of Minnesota Press.
- Goldberg, L. R. (1965). Diagnosticians versus diagnostic signs: The diagnosis of psychosis versus neurosis from the MMPI. *Psychological Monographs, 79*, -(9, Whole No. 602).
- Gordon, R. M. (2001). MMPI/MMPI-2 changes in long-term psychoanalytic

- psychotherapy. *Issues in Psychoanalytic Psychology*. Vol 23(1-2) , 59-79.
- Gordon, R. M. (2006). False assumptions about psychopathology, hysteria and the MMPI-2 restructured clinical scales. *Psychological Reports*, 98, 870-872.
- Gordon, R. M. (2007, spring). The powerful combination of the MMPI-2 and the Psychodynamic Diagnostic Manual, *Independent Practitioner*, 84–85.
- Gordon, R. M., Stoffey, R., & Bottinelli, J. (2008). MMPI-2 findings of primitive defenses in alienating parents. *American Journal of Family Therapy*, 36 (3): 211–228.
- Gordon, R.M. and Stoffey, R.W. and Perkins, B.L. (2013) Comparing the Sensitivity of the MMPI-2 Clinical Scales and the MMPI-RC Scales to Clients Rated as Psychotic, Borderline or Neurotic on the Psychodiagnostic Chart, *Psychology: Special issue on Criminal Investigative Psychology*, 4, 9A, 12-16. doi: 10.4236/psych.2013.49A1003.
- Graham, J. R., Ben-Porath, Y. S., & McNulty, J. (1999). Using the MMPI-2 in outpatient mental health settings. Minneapolis: University of Minnesota Press.
- Gynther, M. D. (1979). Aging and personality. In James N. Butcher (Ed.), *New developments in the use of the MMPI*. Minneapolis, MN: University of Minnesota Press.
- Hopwood, C.J. & Bornstein, R.F. (Eds) (2014) *Multimethod Clinical Assessment*. The Guildford Press, New York.
- McKinley, J. C., Hathaway, S. R., & Meehl, P. E. (1948). The Minnesota Multiphasic Personality Inventory: VI. The K scale. *Journal of Consulting Psychology*, 12, 20-31.
- Spiro III, A., Butcher, J.N., Levenson, R.M., Aldwin, C.M., and Bosse, R. (2000). Change and Stability in Personality: A Five Year Study of the MMPI-2 in Older Men. Butcher, J. E. *Basic Sources on the MMPI-2*. (443-462). Minneapolis: University of Minnesota Press.
- Tellegen, A., Ben-Porath, Y.S., McNulty, J.L., Arbisi, P.A., Graham, J.R., & Kaemmer, B. (2003). *The MMPI-2 Restructured Clinical Scales: Development, validation, and interpretation*. Minneapolis, MN: University of Minnesota Press.

The Personality Assessment Inventory (PAI)

The PAI is a 344-item self-report measure of psychopathology, developed to provide measures of constructs that are central in treatment planning, implementation, and evaluation (Morey, 1991, 1996). Items are rated on a 4-point Likert scale (from “false” to “very true”). There is an adolescent version as well. Both require a 4th grade reading level.

Twenty-two nonoverlapping scales measure a broad range of psychological constructs: four validity scales (Inconsistency, Infrequency, Negative Impression Management and Positive Impression Management), 11 clinical scales (Somatization, Anxiety, Anxiety-Related Disorders, Depression, Mania, Paranoia, Schizophrenia, Borderline Features, Antisocial Features, Alcohol Use, and Drug Use), 5 treatment consideration scales (Aggression, Suicidal Ideation, Stress, Nonsupport, and Treatment Rejection), and 2 interpersonal scales (Dominance and Warmth). Ten of the full scales contain “conceptually derived subscales (p.3, Morey, 1996).”

There are also supplemental indices, some of which include Mean Clinical Elevation (general psychiatric distress), suicide and violence potential, and a treatment process index (composite score predicting treatment amenability).

Computer scoring offers an interpretative report and diagnostic considerations. Normative data include 1,000 nonpatients age-distributed according to census data, a large mixed sample of inpatients and outpatients, and a large college student sample. Results are profiled as T-Scores relative to the census sample. The PAI profile includes indications of unusual elevations on the basis of data from the reference sample of patients. Extensive reliability and validity data are available in the literature, much of which is summarized in the revised edition of the test manual (Morey, 2007) and in two other interpretive and summarizing books (Morey, 1996; Blais, Baity, & Hopwood, 2010)

Strengths. The PAI is easy to administer and requires a relatively low reading level. It “differs from other well-known self-report multi-scale inventories in several important ways that are largely a consequence of the construct validation approach to test construction (Hopwood, Blais, & Baity, 2010, p. 1).” It has a dimensional response scale

(4-point Likert), rather than a categorical response one (True-False). Items are non-overlapping. There are normative data for specialty populations, including personnel selection and forensics. It has a high degree of internal consistency across samples; results are stable over periods of 2-4 weeks ($r > .80$ for the 22 scales). Validity studies demonstrate convergent and discriminant validity with more than 50 other measures of psychopathology. It includes a short form, calculating 20 of the 22 full scales, that is based on the first 160 items for patients who fatigue easily or if the clinician wishes a quick overview. The short form has adequate to strong psychometric properties (Morey, 1991; Siefert et al., 2012; Sinclair et al., 2010; Sinclair et al., 2009).

The PAI assesses both state (i.e., depression, anxiety, mania, etc.) and trait (i.e., borderline and antisocial features) phenomena, useful in assessing personality features (P Axis) and overt or manifest symptoms (S Axis). The PAI can also assist in the assessment of M Axis (Blais & Hopwood, 2010).

Limitations. The PAI is a widely used broadband measure of psychological functioning, but as with any self-report measure, it is based purely on patient's perspective. Also, because of its breadth and depth it is long, even in its shorter version.

For more information, write to Michelle Stein (MSTEIN3@mg.harvard.edu).

References

- Blais, M.A., Baity, M.R., & Hopwood, C.J. (2010). *Clinical Applications of the Personality Assessment Inventory*. New York, NY: Routledge.
- Blais, M.A. & Hopwood, C.J. (2010). *Personality Focused Assessment with the PAI*. In Blais, M.A., Baity, M.R., & Hopwood, C.J. (Eds.) *Clinical Applications of the Personality Assessment Inventory*. New York, NY: Routledge.
- Hopwood, C.J., Blais, M.A., & Baity, M.R. (2010). *Introduction* In Blais, M.A., Baity, M.R., & Hopwood, C.J. (Eds.) *Clinical Applications of the Personality Assessment Inventory*. New York, NY: Routledge.

- Morey, L.C. & Ambwani, S. (2008). *The Personality Assessment Inventory* In Boyle, G., Matthews, G. & Saklofske, D.H. (Eds. Vol. 2) *The Sage Handbook for Personality Theory and Assessment*. London: Sage.
- Morey, L.C. & Hopwood, .J. (2007) *Casebook for the Personality Assessment Inventory*. Odessa, FL: Psychological Assessment Resources
- Morey, L.C. (2007). *The Personality Assessment Inventory professional manual*. Lutz, FL: Psychological Assessment Resources.
- Morey, L.C. (1991/1996). *An Interpretive Guide to the Personality Assessment Inventory (PAI)*: Odessa, FL: Psychological Assessment Resources.
- Siefert, C.J., Stein.M, Sinclair, SJ, Antonius, D, Shiva, A, & Blais, M.A.(2012). Development and Initial Validation of a Scale for Detecting Inconsistent Responding on the Personality Assessment Inventory-Short Form. *Journal of Personality Assessment*, 94(6), 601-606.
- Sinclair, S. J., Antonius, D., Shiva, A., Siefert, C. J., Kehl-Fie, K., Lama, S., . . . Blais, M. A. (2010). The psychometric properties of the Personality Assessment Inventory–Short Form (PAI–SF) in inpatient forensic and civil samples. *Journal of Psychopathology and Behavioral Assessment*, 32, 406–415.
- Sinclair, S. J., Siefert, C. J., Shorey, H., Antonius, D., Shiva, A., Kehl-Fie, K., & Blais, M. A. (2009). A psychometric evaluation of the Personality Assessment Inventory–Short Form (PAI–SF) clinical scales in an inpatient psychiatric sample. *Psychiatry Research*, 170, 262–266.

The Millon Clinical Multiaxial Inventory-III (MCMI-III)

The MCMI is anchored in Theodore Millon’s (2011) evolutionary theory of personality and psychopathology. The MCMI was designed to assess both Axis I and Axis II DSM disorders and to assist clinicians in formulating diagnoses and in developing a treatment plan that takes into account the patient’s personality style and coping behavior. The current version, the MCMI-III was published in 1994 and reflected the revisions in DSM-IV. It has 175 true-false questions and usually takes about 30 minutes to complete. It can be used by

clients 18 years old or older and requires an 8th grade reading level. It also now uses combined gender norms.

The MCMI-III assesses 15 personality styles and subtypes: Retiring/Schizoid, Shy/Avoidant, Pessimistic/Melancholic, Cooperative/Dependent, Exuberant/Hypomanic, Sociable/Histrionic, Confident/Narcissistic, Nonconforming/Antisocial, Assertive/Sadistic, Conscientious/Compulsive, Skeptical/Negativistic, Aggrieved/Masochistic, Eccentric/Schizotypal, Capricious/Borderline, and Suspicious/Paranoid.

The 10 Clinical Syndrome Scales (which coincide with DSM-IV Axis I disorders) are: Anxiety, Somatoform, Bipolar Manic, Dysthymia, Alcohol Dependence, Drug Dependence, Posttraumatic Stress Disorder, Thought Disorder, Major Depression, and Delusional Disorder. There are 5 Validity and Bias Scales.

After scales were rationally constructed, which means that items have face validity and it is a relatively easy test to fake good on, individuals diagnosed with each of the disorders in the scale were administered the test, and cut-offs for the various levels of each scale were determined in accordance with the number of individuals in the pool with the particular disorder. The MCMI is the second most frequently used instrument after the MMPI, and this popularity might be due in part to the quality of the computer-generated narrative descriptions that are readable, clinically applicable, and descriptive of the patient in plain English. Another asset is the use not of conventional percentile or standardized scores but of Base Rate scores (BRs) that incorporate prior probabilities of each disorder in test output and diagnostic thresholds.

The MCMI has been adapted for assessment of adolescents as the Millon Adolescent Clinical Inventory (MACI). Construction of the MACI has involved changes in items, age appropriate norms, and name revisions for the scales. The interpretive strategies are different as well. For instance, there is more emphasis on interpreting individual scales than configurations of scales. Similarly, the MCMI has been adapted for pre-adolescents as the Millon Pre-Adolescent Clinical Inventory (M-PACI).

Strengths. The MCMI, MACI, and M-PACI can be useful aids in the assessment of virtually all PDM-2 axes for adults, adolescents and children. One of the strengths of the MCMI is that it is normed on a pathological population, and the BR scores indicate the probability or confidence of the diagnosis. The MCMI-III has good validity and reliability: Cronbach's alpha statistics range from .66 (Compulsive) to .90 (Major Depression). In a sample of 87 participants, the test-retest reliability of the MCMI-III (5–14 days later) ranged from .82 (Debasement) to .96 (Somatoform), with a median coefficient of .91. Positive predictive power ranged from .30 (Masochistic) to .81 (Dependent). Sensitivity, or the proportion of individuals that have a condition that are correctly identified ranged from .44 (Negativistic) to .92 (Paranoid). The ability of some of the self-report scales – borderline, for instance – to identify personality reliably pathology is impressive.

Limitations. The MCMI-III is highly stable over a short period of time; however, no long-term data are available. As with the MMPI-2 and the PAI, the MCMI-III is a self-report measure and can be affected by lack of insight, and it does not have as many or as well-researched validity scales as the MMPI-2. The MCMI-III is based on clinical samples and is applicable only to individuals who evidence problematic emotional and interpersonal symptoms or who are undergoing treatment or a psychodiagnostic evaluation.

References

- McCann, J. T. (1999). *Assessing adolescents with the MACI: Using the Millon Adolescent Clinical Inventory*. John Wiley & Sons Inc.
- Millon, T. (2011). *Disorders of personality: introducing a DSM/ICD spectrum from normal to abnormal (Vol. 208)*. John Wiley & Sons.
- Millon, T., Millon, C., Davis, R., & Grossman, S. (2006). *MCMI-III Manual (Third ed.)*. Minneapolis, MN: Pearson Education, Inc.

The Severity Indices of Personality Problems-118

The SIPP-118 (Verheul et al., 2008) is a self-report questionnaire covering core

components of personality functioning. The assumption upon which the SIPP is founded is that personality pathology can be understood as a consequence of alterations of human adaptive capacities, and the aim of the SIPP-118 is to assess these capacities.

The SIPP contains 118 items, each answered on a 4-point, Likert scale from 1 (fully disagree) to 4 (fully agree), for 16 domains: emotion regulation, effortful control, stable self-image, self-reflective functioning, aggression regulation, frustration tolerance, self-respect, purposefulness, enjoyment, feeling recognized, intimacy, enduring relationships, responsible industry, trustworthiness, respect, and cooperation.

These domains constitute homogeneous item clusters (i.e., the clusters are unidimensional and internally consistent) that fit well into 5 clinically relevant higher-order domains: self-control, identity integration, relational capacities, social concordance, and responsibility. These domains appeared to have good concurrent validity across various populations, good convergent validity in terms of associations with interview ratings of the severity of personality pathology, and good discriminant validity in terms of associations with trait-based personality disorder dimensions.

The domain scores are stable over a time interval of 14–21 days in a student sample but are sensitive to change over a 2-year follow-up interval in a treated patient population. The SIPP-118 provides a set of five reliable, valid, and efficient indices of the core components of (mal)adaptive personality functioning. Initial research has shown a cross-national validity of this tool (Arnevik et al., 2009).

Strengths. The SIPP can be a useful aid for the assessment of basic mental capacities of both adolescent and adult subjects (M and MA Axes). It is a clinically relevant, psychometrically sound, and user-friendly tool. Available in seven different languages, it can be used both for clinical assessment and for research in psychotherapy.

Limitations. The SIPP-118 does not provide a personality or clinical diagnosis of a patient. Research conducted with this tool thus far has been limited. The instrument is less well known and used than other self-report instruments reviewed in this section. To date, its degree of overlap with other interpersonal measures, such as the IIP (see below), is

unknown.

For more information, see www.deviersprong.nl/paginas/143-sipp-main-menu.html.

References

- Verheul, R., Berghout, C.C., Busschbach, J.J.V., Bateman, A., Helene, A., Dolan, C., van der Kroft, P.J.A., & Fonagy, P. (2008). Severity Indices of Personality Problems (SIPP-118): Development, Factor Structure, Reliability, and Validity. *Psychological Assessment*, 20, 2, 23-34.
- Arnevik, E., Wilberg, T., Monsen, J.T., Andrea, H., & Karterud, S. (2009). A cross-national validity study of the Severity Indices of Personality Problems (SIPP-118). *Personality and Mental Health*, 3, 41-55.
- Feenstra, D.J., Hutsebaut, J., Verheul, R., & Busschbach, J.J.V. (2011). Severity Index of Personality Problems (SIPP-118) in adolescents: reliability and validity. *Psychological Assessment*, 23, 646-655.

The Inventory of Interpersonal Problems (IIP)

The IIP is a self-report inventory designed to identify problematic areas of relational functioning (Horowitz, Alden, Wiggins, & Pincus, 2000; Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988). It is based on a circumplex model of interpersonal behavior that draws from a rich theoretical and empirical literature (e.g., Alden et al., 1990; Benjamin, 1974; Kiesler, 1983; Leary, 1957; Wiggins, 1979, 1991). This model is represented by a two-dimensional circumplex defined by two orthogonal and bipolar axes of interpersonal style: dominance vs. submission and love vs. hate. The circumplex space is divided into eight octants that form a circular array of blends of the dominance and love dimensions and can reflect various interpersonal problems.

The original IIP (Horowitz et al., 1988) consisted of 127 items, but the most widely used version of this instrument is the Inventory of Interpersonal Problems – Circumplex (IIP-C; Alden et al., 1990; Horowitz et al., 2000), comprising 64 items. Many additional

short forms (see, for example, the IIP-32; Barkham, Hardy, & Startup, 1996) and derivative tools containing item sets have been developed to accomplish specific purposes (e.g., screen for personality disorders). Also various methods for scoring the IIP are available (Hughes & Barkham, 2005; Gurtman, 2006). In this section we focus on the IIP-C.

The 64 items of the IIP-C assess the interpersonal behaviors that an individual shows in excess (e.g., “I fight with other people too much”) or areas of difficulty (e.g., “It is hard for me to join in groups”). Each item is rated on a 7-point Likert scale ranging from 1 (Not at All) to 7 (Extremely). The overall score of the IIP-C is used as an indicator of general interpersonal problems, but the IIP-C also yields scores on eight subscales indicating specific difficulties: domineering, vindictive, cold, socially inhibited, nonassertive, exploitable, overly nurturant, and intrusive.

The IIP-C has strong psychometric properties. Internal consistency alpha coefficients for the subscales range from .76 to .88, and test-retest reliabilities for scales range from .58 to .84 (Horowitz et al., 2000). Subscales of the IIP-C converge with measures tapping similar constructs in clinical (Gurtman, 2006; Haggerty, Hilsenroth, Vala-Stewart, 2009) and nonclinical samples (Alden et al., 1990; Horowitz et al., 2000). This measure is also sensitive to change and, for this reason, is frequently employed in psychotherapy outcome research (e.g., Ruiz, Pincus, Borkovec, Echemendia, Castonguay, & Ragusea, 2004).

Strengths. The IIP may be a useful aid in the assessment of some personality features (P Axis) and of the capacity for relationships and intimacy of the M Axis. It is easy to administer and requires minimal training. It has good psychometric properties and, in particular, is reliable over time, sensitive to change, convergent with similar tools, and predictive of treatment outcome. It can be effectively employed with clinical and nonclinical samples. Finally, it provides an overall assessment of interpersonal functioning, as well as of specific relational domains.

Limitations. The IIP-C suffers from the same biases as other self-report measures: it can

be faked and is affected by the patient's limitations in acknowledging, mentalizing or communicating difficulties. Moreover, this measure does not include any validity scales. The normative data and clinical cut-offs are not clearly established on large, representative samples.

For more information, see: www.mindgarden.com/products/iip.htm.

References

- Alden, L.E., Wiggins, J.S., & Pincus, A.L. (1990). Construction of circumplex scales for the Inventory of Interpersonal Problems. *Journal of Personality Assessment*, *55*, 521-536.
- Gurtman, M.B. (2006). Interpersonal problems and the psychotherapy context: The construct validity of the inventory of interpersonal problems. *Psychological Assessment*, *8*, 241-255.
- Horowitz, L.M., Alden, L.E., Wiggins, J.S., & Pincus, A.L. (2000). *Inventory of Interpersonal Problems manual*. San Antonio, TX: The Psychological Corporation.
- Horowitz, L.M., Rosenberg, S.E., Baer, B.A., Ureño, G., & Villaseñor, V.S. (1988). Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, *56*, 885-892.

The Central Relationship Questionnaire (CRQ)

In 1998, Barber, Foltz, and Weinryb constructed the Central Relationship Questionnaire (CRQ) to measure interpersonal patterns. In 2009, revisions were made to the CRQ. The 101-item revised Central Relationship Questionnaire-Revised (CRQ-R; McCarthy, Connolly Gibbons, & Barber, 2008) is a self-report measure of the Core Conflictual Relationship Theme (CCRT described above; Luborsky & Crits-Christoph, 1998). It involves revisions of the original CRQ to increase the interpersonal dimensions it captures, reduce its length, and model a higher-order factor structure (McCarthy et al., 2008).

The CRQ-R is used to assess representations of the patients' wishes in interpersonal relationships, their perceptions of the responses of others to their wishes, and their own responses to the other. Participants are instructed to rate each of their four central relationships (relationships with one's romantic partner, mother, father, and best friend) in terms of the three main relationship themes: Wishes (W), Responses from Other (RO), and Responses of Self (RS). Participants rate the likelihood that a particular interpersonal theme is present in each of their central relationships on a 7-point Likert scale. Scores for each interpersonal theme can be aggregated across relationships, such that each patient will have one score on each of the 16 themes representing the patient's general representations of interpersonal relationships: 5 types of Ws (be independent, be intimate, be hurtful, be sexual, be submissive), 5 types of ROs (is hurtful, is independent, is loving, is sexual, is submissive), and 6 types of RSs (am autonomous, am avoidant, am domineering, am intimate, am nonconfrontational, am sexual).

Strengths. The CRQ can be an useful aid for assessing object relationships on both the P Axis and the M axis. The psychometric properties of the revised CRQ were found to be adequate (Barber, Foltz, & Weinryb, 1998). Several studies have demonstrated the relevance and contributions of the CRQ to psychotherapy research. For example, greater rigidity across interpersonal relationships (as measured with the CRQ), was related to fewer symptoms and interpersonal problems (McCarthy et al., 2008). Moreover, patients' pretreatment representations of significant others (as measured by the CRQ) predicted a substantial part of the therapeutic alliance throughout the course of treatment (Zilcha-Mano, McCarthy, Dinger, & Barber, 2014).

Limitations. Although the CRQ-R is shorter than the original CRQ, it is still time consuming, especially when the patient needs to rate 101 items for each of several significant others in his or her life. The instrument has not received much research or clinical use beyond the developers.

For more information, write to Kevin Scott McCarthy (kmccarth@psych.upenn.edu).

References

- Barber, J. P., Foltz, C., & Weinryb, R. M. (1998). The Central Relationship Questionnaire: Initial report. *Journal of Counseling Psychology, 45*, 131–142.
- McCarthy, K. S., Connolly Gibbons, M. B., & Barber, J. P. (2008). The relation of consistency in interpersonal patterns to symptoms and functioning: An investigation using the Central Relationship Questionnaire. *Journal of Counseling Psychology, 55*, 346-358.
- Zilcha-Mano, S., McCarthy, K. S., Dinger, U., & Barber, J. P. (2014). To what extent is the alliance affected by transference? Some empirical findings. *Psychotherapy, 51*, 424-433. <http://dx.doi.org/10.1037/a0036566>
- Zilcha-Mano, S., Chui, H., McCarthy, K., Dinger, D., & Barber, J. P. (in preparation). Examining mechanisms underlying the dodo bird effect in depression: changes in general attributions and relationship representations.

The Toronto Alexithymia Scale (TAS-20)

Alexithymia is a multifaceted personality construct representing a deficit in the cognitive processing of emotions. It is composed of two higher order factors including deficit of affect awareness (difficulty identifying and describing feelings) and operatory thinking (externally oriented thinking and poor imaginal processes) (Taylor & Bagby, 2012). It is strongly influenced by early interactions with caregivers because inadequate responses to the child's emotions have a major influence on the ability to self-regulate both emotional and neurobiological states later in adulthood (Taylor, Bagby, & Parker, 1997). The most widely used assessment instrument is the 20-item version of the Toronto Alexithymia Scale (TAS-20; Bagby, Parker, & Taylor, 1994a, 1994b), a self-report questionnaire that assesses three facets of alexithymia: difficulty identifying feelings, difficulty describing feelings, and externally oriented thinking. Items are rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Although alexithymia

is a dimensional construct, an empirically derived score of >60 identifies subjects in the higher alexithymia range.

Strengths. The TAS-20 can be an useful aid for assessing affect regulation, deficits in mentalizing affective states, inability to use symbolic expressions, and emotional arousal dissociated from psychological meanings (Taylor, 2010) both on the P Axis and on the M Axis. Extensive validation, replication of the factor structure in several languages and countries, short administration time, and ease of use have been among the reasons why the TAS-20 has become the reference standard for measuring alexithymia in several psychiatric and medical settings (Lumley, Neely, & Burger, 2007).

Limitations. There are doubts that alexithymia can be validly assessed by a self-report scale because it is, by definition, the inability to report individual psychological states. Also, implicit dimensions of personality related to skills (as intelligence) or lack of abilities (as alexithymia) are more reliably assessed with non-self-report measures, such as performance-based or observer-based instruments. In fact, self-report measures of alexithymia and related constructs showed good correlations with each other (r in the .40-.60 range) but lower associations with different methods (interviews, collateral reports, performance testing) (r in the .20 –.30 range) for the same constructs (Lumley, Gustavson, Partridge, & Labouvie-Vief, 2005). Therefore, the need for measures of alexithymia based upon clinical observation has been repeatedly underscored in recent years. These include the Observer Alexithymia Scale (OAS) (Haviland, Warren, & Riggs, 2000), the Toronto Structured Interview for Alexithymia (TSIA) (Bagby, Taylor, Parker, & Dickens, 2006), the Diagnostic Criteria for Psychosomatic Research (DCPR)-Alexithymia cluster (Porcelli & De Carne, 2001), and the Rorschach Alexithymia Scale (RAS (Porcelli & Mihura, 2010).

For more information, you can write to Piero Porcelli (piero.porcelli@ircsdebellis.it)

References

- Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994a). The twenty-item Toronto Alexithymia Scale: I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*, 23–32.
- Bagby, R.M., Parker, J. D. A., & Taylor, G. J. (1994b). The twenty-item Toronto Alexithymia Scale: II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research, 38*, 33–40.
- Bagby, R. M., Taylor, G. J., Parker, J. D., & Dickens, S. E. (2006). The development of the Toronto Structured Interview for Alexithymia: Item selection, factor structure, reliability and concurrent validity. *Psychotherapy and Psychosomatics, 75*, 25–39.
- Haviland, M. G., Warren, W. L., & Riggs, M. L. (2000). An observer scale to measure alexithymia. *Psychosomatics, 41*, 385–392.
- Lumley, M. A., Gustavson, B. J., Partridge, R. T., & Labouvie-Vief, G. (2005). Assessing alexithymia and related emotional ability constructs using multiple methods: Interrelationships among measures. *Emotion, 5*, 329–342.
- Lumley, M. A., Neely, L. C., & Burger, A. J. (2007). The assessment of alexithymia in medical settings: Implications for understanding and treating health problems. *Journal of Personality Assessment, 89*, 230–246.
- Porcelli, P., & De Carne, M. (2001). Criterion-related validity of the Diagnostic Criteria for Psychosomatic Research for alexithymia in patients with functional gastrointestinal disorders. *Psychotherapy and Psychosomatics, 70*, 184–188.
- Porcelli, P., & Mihura, J.L. (2010). Assessment of alexithymia with the Rorschach Comprehensive System: the Rorschach Alexithymia Scale (RAS). *Journal of Personality Assessment, 92*, 128-136.
- Taylor, G.J. (2010). Affects, trauma, and mechanisms of symptom formation: A tribute to John C. Nemiah, MD (1918–2009). *Psychotherapy and Psychosomatics, 79*, 339-349.
- Taylor, G.J., & Bagby, R.M. (2012). The alexithymia personality dimension. In T.A. Widiger (Ed.), *The Oxford handbook of personality disorders* (pp.648-676). New York: Oxford University Press.

Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (1997). *Disorders of affect regulation*. Cambridge, UK: Cambridge University Press.

The Clinical Outcome in Routine Evaluation (CORE)

This self-report measure, developed by Chris Evans and colleagues in Britain to serve the broad needs of the British National Health Service (NHS) was developed from a pan-theoretical perspective. Thirty-four items address the previous seven days, covering domains of well-being (e.g. “Over the last week ... I have felt O.K. about myself”), problems (“I have felt panic or terror”), functioning (“I have felt able to cope when things go wrong”) and risk (internal, “I have thought of hurting myself,” and external, “I have been physically violent to others”) are rated on a 5 point frequency scale from 0 (Not at All) to 5 (Most or All of the Time). Eight items are reverse scored to ensure that there is some variety across the items.

An extensive survey of psychotherapists and counselors of different modalities, different core professions (e.g., psychology, psychiatry, social work), and different employment settings identified core items. Initial psychometric exploration showed acceptability and usability, and good internal and test-retest reliability (.94 and .90 respectively, Cronbach's alpha and Spearman rho), small effects of age and gender, good convergent validity with established self-report measures of depression, good sensitivity to change and to discrimination between clinical and nonclinical groups (Cohen *d*, 1.65: Evans et al. 2002). These psychometric properties have been replicated in clinical samples in the original English and in an increasing number of translations into other languages. Reliable change indicators like those of Jacobson and Truax (1991) and clinically significant change cutting points were given and are becoming available for translations.

The self-report measures were complemented by a practitioner-completed CORE-A (Assessment) comprising the Therapy Assessment Form (TAF) and the End of Therapy (EoT) form. The TAF asked about social situation, focal problems, past mental health care, medication, and diagnosis (optional) and about the therapy planned (TAF). The EoT covered details about the therapy actually given.

A short form was designed to be used in general population, nonclinical survey work (CORE-GP). and even shorter form, the CORE-10 was produced, and there are a variety of other versions for special purposes.

Recent developments include a scoring method that converts CORE-OM scores (using scores on six of the items) to health economic “quality equivalence,” making possible conversion from CORE-OM gains to QALY (Quality Adjusted Life Years) gains (Mavranouzouli et al. 2011; Mavranouzouli et al. 2012), and SCORE (Systemic CORE): a self-report evaluation of families (not the individuals in them) for use in family-systemic therapies and based on CORE system principles (Stratton et al. 2010; Stratton et al. 2013). A summary of the CORE system is provided in Evans (2012) and is free to download, as is all the information about the system at www.coresystemtrust.org.uk/instruments/core-5-information/.

Strengths. All CORE instruments are completely free. The CORE-OM has shown good psychometric properties in all evaluations, even in translation, and has been used with a very wide diversity of therapies. Shortened measures exist and, like the CORE-OM, these have been translated into 25 languages to a high quality translation protocol with the number of translations rising steadily. Adaptations and extensions have been constructed for people with learning disabilities, young people and for evaluation of families. The website also hosts a growing collection of information about the system and an index of publications using CORE instruments or elements of the CORE system.

Limitations. No measures or system can ever provide a complete picture of change in therapy, and all self-report measures are open to manipulation.

For more information contact (<https://www.coresystemtrust.org.uk/>).

References

Evans, C., 2012. The CORE-OM (Clinical Outcomes in Routine Evaluation) and its

derivatives. *Integrating Science and Practice*, 2(2). Available at:

http://www.ordrepsy.qc.ca/pdf/2012_11_01_Integrating_SandP_Dossier_02_Evans_En.pdf.

- Evans, C. et al., 2002. Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180(JAN.), pp.51–60.
- Jacobson, N.S. & Truax, P. (1991). Clinical Significance: a Statistical Approach to Defining Meaningful Change in Psychotherapy Research.. *Journal of Consulting and Clinical Psychology* 59, 12-19.
- Mavranouzouli, I. et al., 2012. Estimating a Preference-Based Index from the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM): valuation of CORE-6D. *Medical Decision Making*, 20, pp.321–333.
- Mavranouzouli, I. et al., 2011. Using Rasch analysis to form plausible health states amenable to valuation: the development of CORE-6D from a measure of common mental health problems (CORE-OM). *Quality of Life Research*, 20(3), pp.321–333.
- Stratton, P. et al., 2013. Detecting therapeutic improvement early in therapy: validation of the SCORE-15 index of family functioning and change: Validation of the SCORE-15 index. *Journal of Family Therapy*, 36(1), pp.3–19.
- Stratton, P. et al., 2010. Developing an indicator of family function and a practicable outcome measure for systemic family and couple therapy: The SCORE. *Journal of Family Therapy*, 32(3), pp.232–258.
- Twigg, E. et al., 2010. The Young Person's CORE: Development of a brief outcome measure for young people. *Counselling and Psychotherapy Research: Linking research with practice*, 9, pp.160 – 168.

The Outcome Questionnaire-45 (OQ-45)

The OQ-45 (Lambert et al., 1996) is a 45-item self-report scale used to estimate client degree of disturbance at the outset and over the course of treatment. It provides an index of mental and interpersonal health or well-being for adults 18 years of age or older. A person

who takes the measure is compared to inpatient, community mental health, outpatient, employee assistance program, college counseling center, and normal populations. Scores on the measure are referenced against expected treatment responses based on the progress of 11,000 treated individuals across the United States. These data provide a benchmark of success in order to identify treatment non-responders and clients at risk for negative outcomes. It also provides cut-scores for reliable change and recovery as markers for gauging treatment success and possible termination of services. It has been translated into more than 30 languages other than English. It requires reading ability at the 6th-grade level.

Because almost all adults who enter treatment experience symptoms of anxiety and depression, half of the items of the OQ-45 measure core aspects of symptomatic distress or subjective discomfort. Because a satisfactory quality of life and well-being depends on positive interpersonal functioning, a quarter of the items assess disturbance in interpersonal relationships with intimate others. The final quarter of the items assess functioning in social roles such as work, school, homemaking, and leisure activities.

The OQ-45 can be used regardless of the type of psychotherapy, mode of psychotherapy, or medication intervention. It is a-theoretical in nature and serves as a mental health vital sign or lab test to be used by clinicians to manage illness by quantifying the patient's current mental health functioning. In contrast to purely symptom measures, the OQ-45 attempts to assess factors that are particularly important to clinicians and patients with a psychodynamic orientation—interpersonal relations and social role functioning.

Nowadays, and ideally, the OQ-45 is administered online, via handheld devices, or a personal computer (it can be administered and scored via hard copy as well). It takes about 5-10 minutes of patient time to rate all of the 45 items, typically prior to the treatment session. Each item (e.g., "I feel hopeless about the future") is answered on a 5 point scale according to the patient's recollection of the preceding week—on a scale from "almost always" to "never." Software (OQ-Analyst) scores the measure, graphs the results in relation to earlier administrations, and in relation to normative functioning and expected treatment response expected from other individuals who have the same initial level of

disturbance. All this is accomplished and available on the therapist's computer immediately after completion of the final item. A clinician report and client report can be printed and delivered to individuals as a hard copy. Cumulative data from multiple administration of the measure can be housed on a self-supported server (or personal computer).

The OQ-45 is part of a larger Outcome Measurement System--the OQ-Analyst, which includes measures of child functioning, the Brief Psychiatric Rating Scale, and the Assessment for Signal Clients, a clinical support tool used to guide problem-solving with failing cases. In addition, OQ-reports estimate the number of sessions needed for a patient to return to a state of normal functioning or at least achieve reliable change.

The OQ-Analyst can be set up to provide clinical information to interdisciplinary teams who are also working with each specific patient. This shared information allows for all the team to be aware of an absent or negative treatment response and settle on coordinated efforts to turn the course of treatment in a positive direction. The OQ-Analyst also produces reports, summed across all patients or patient subgroups (such as substance-abuse patients, males vs. females, program A versus program B) for the purpose of comparing benchmarks across similar services or for particular clinicians.

Dozens of studies have been published on the psychometric properties of the OQ-45. It has high internal consistency (.90), test-retest reliability (.84 over 3 weeks), and concurrent validity with scales such as the SCL-90 and Beck Depression Inventory, with coefficients hovering in the mid .80s. Factor analytic studies support the presence of an overall distress factor with three subordinate factors consistent with the subscales. Most items, the subscales, and the Total Score are sensitive to the effects of interventions while remaining stable in untreated individuals.

Strengths. The OQ-45 can be a useful aid for the assessment of M Axis overall level of mental functioning, in particular of the level of subjective well-being and adjustment experienced by the patient . It assesses general psychological distress, interpersonal functioning, and role functioning. Versions are available for youths, adults, and even

patients in group therapy.

The OQ-45 is supported by extensive published evidence of the degree to which providing feedback to clinicians and patients based on the OQ-45 alert system maximizes patient outcomes and reduces treatment failures. Eleven RCT's have been completed showing that the feedback and problem-solving tools delivered to therapists work in a variety of routine care settings. The OQ-45 has been judged by the National Registry of Evidence-based Programs & Practices (NREPP) as an evidence-based practice. NREPP rated the degree to which the OQ-Analyst is ready for dissemination at 3.9 on a 4-point scale. Qualitative research has been completed indicating that clients do not mind taking the OQ-45 on a weekly basis. Providers are less positive initially than clients, especially if use of the measure is forced on them by administrators, but after a time even the most resistant clinicians come to appreciate the feedback graphs of patient progress.

Limitations. The OQ-45 is a self-report measure that suffers from the same limitations as other self-report measures. It is easy to manipulate the scores one way or another if the patient is motivated to do so. It can be used by patients to test what therapists will do if they answer in a particular way, such as exaggerating their scores to punish a therapist. The instrument and the scoring are expensive and typically require computer use. In addition, because the measure was not developed from psychodynamic theories, it has not shed light on constructs of special importance to psychodynamic psychotherapists, and neither does it lend itself to the task of diagnosis. Most of the research supporting its use is based on short-term treatments lasting no more than 20 sessions.

For more information, see www.oqmeasures.com/.

References

- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. M., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying patients at risk for treatment failure. *Journal of Clinical Psychology: In Session*, 61(2), 155-163.
- Lambert, M. J. (2010). Prevention of treatment failure: The use of measuring, monitoring,

& feedback in clinical practice. Washington, DC: APA Press.

Lambert, M. J., Hansen, N. B., & Harmon, S. C. (2010). The OQ-45 system: Development and practical applications in health care settings. In M. Barkham, G. Hardy, & J. Mellor-Clark (Eds.) *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 141-154). New York: Wiley-Blackwell.

Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology and Psychotherapy*, 3, 249–258

Lambert, M. J., Kahler, M., Harmon, C., Burlingame, G. M., & Shimokawa, K. (2013). *Administration & scoring manual for the Outcome Questionnaire-45.2*. Salt Lake City: OQMeasures.

Okiishi, J., Lambert, M. J., Eggett, D., Nielsen, S. L., Dayton, D. D., & Vermeersch, D. A. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists on their clients' psychotherapy outcome; *Journal of Clinical Psychology*, 62, 1157-1172.

Shimokawa, K., Lambert, M.J., & Smart, D. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic & mega-analytic review of a psychotherapy quality assurance program. *Journal of Consulting & Clinical Psychology*, 78, 298-311.

Vermeersch, D A., Whipple, J. L., Lambert, M. J., Hawkins, E. J., Burchfield, C. M., & Okiishi, J. C. (2004). Outcome Questionnaire: Item sensitivity to changes in counseling center clients. *Journal of Counseling Psychology*, 51, 38-49.

The Schwartz Outcome Scale 10 (SOS-10)

The SOS-10 (Blais et al., 1999) is a distinctive, low-burden measure developed to monitor outcomes, at both the individual and group level. Construction of the SOS-10 was guided by insights obtained from a diverse group of senior clinicians and patients.

Interviews conducted with senior psychologists, psychiatrists, and a neurosurgeon, as well as with patient focus groups, were used to discover the changes that occurred (excluding

symptoms) with successful treatment. The interviews and focus group discussions were transcribed and reviewed for common themes. Common themes were used to generate an initial item pool. Empirical evaluation and refinement identified 20 well-performing items, and Rasch analysis was employed to reduce the scale to its final 10-item version (Blais, et al., 1999). Initially developed for use with adults (age 17 and up) recent research has extended the utility of the SOS to adolescent populations.

The SOS-10 has 10 items rated on a 0 (Never) to 6 (All or Nearly All the Time) scale. Higher scores represent greater psychological health and well-being, and lower scores indicate emotional distress and poorer psychological health. The SOS-10 can be administered in traditional paper-and-pencil format or electronically. It is recommended that patients complete the scale prior to a treatment appointment. This way the clinician can determine whether the SOS-10 was completed and review the total score for clinical implications prior to the session.

Owen and Imel (2010) outline a rationale and a practice-friendly procedure for incorporating the SOS-10 into ongoing clinical care. The availability of nonpatient reference data allows for calculation of both a Reliable Change Index and Clinically Significant Improvement (Blais et al., 2011). The ability to apply more sophisticated treatment effectiveness analyses enhances the information obtained from treatment as usual outcome-monitoring programs and increases the comparability of findings across studies.

Scores on the SOS-10 can also be used to identify a patient's level of emotional distress or psychological dysfunction. The following distress ranges, which draw on data from over 10,000 outpatients may prove helpful markers: *Minimal* (59-40), *Mild* (39-33), *Moderate* (32-23) and *Severe* (22-1). Accurately identifying a patient's level of distress at the outset of treatment can help clarify the intensity of services needed, e.g., weekly individual psychotherapy, multiple sessions per week or multiple forms of treatments. Lastly, because SOS-10 items are *not* directly related to psychiatric symptoms, reviewing responses to individual items with patients can afford a nonthreatening avenue for discussing personal strengths and weaknesses

The SOS-10 has solid psychometric properties. Its internal consistency has ranged from

0.84 to 0.96. The test-retest reliability for the scale is also strong, with studies reporting retest correlations of 0.86 and 0.87. Multiple studies both in the original English and in translations have found the SOS-10 to be unifactorial. The accumulated research also supports the construct validity of the SOS-10 as a broad measure of psychological functioning (Blais, et al., 1999; Haggerty et al., 2010 & Young, et al., 2004). The SOS-10 correlates significantly with the Outcomes Questionnaire-45 (OQ-45; Lambert et al., 1996).

Strengths. The SOS-10 may help clinicians to assess patients on the Axes M, MC and MA, in particular the impact of global mental health on patient subjective well-being. It reliably and quickly assesses the conscious distress level of the patient. The SOS-10 has also demonstrated sensitivity to change for a wide variety of treatment modalities, including early treatment change (Hilsenroth, et al., 2001). The SOS-10 has been employed as an outcomes measure in studies of Psychodynamic Psychotherapy, Dialectical Behavior Therapy, residential treatment for refractory Obsessive Compulsive Disorder, inpatient psychiatric treatment as usual and inpatient substance abuse treatment as usual (Blais et al., 2011; Blais et al., 2013).

Limitations. The SOS-10 is a brief self-report instrument with no measures of personality disorders or interpersonal functioning. It suffers from all the response style manipulation concerns associated with such tools, and does not contain validity scales to identify such response styles. . Although the SOS-10 is a proprietary instrument, the scale is made available free of charge for practitioners, researchers, and nonprofit healthcare organizations. While the instrument is easier to complete than the OQ-45 (see above), it does not have the same breadth of development.

For more information write to Mark Blais (MBLAIS@mgh.harvard.edu).

References

Blais, M. A., Lenderking, W. R. Baer, L., deLorell, A., Peets, K., Leahy, L., & Burns, C. (1999).

- Development and initial validation of a brief mental health outcome measure. *Journal of Personality Assessment*, 73, 359-373.
- Blais, MA, Malone, J, Stein, M, Slavin-Mulford, J, Renna, M, & Sinclair, SJ. (2013). Treatment as usual (TAU) for depression: A comparison of psychotherapy, pharmacotherapy and combined treatment at a large academic medical center, *Journal of Psychotherapy*,50: 110-118.
- Blais, M. A., Sinclair, S., Baity, M., Worth, J., Weiss, A., Ball, L., et al. (2011). Measuring outcomes in adult outpatient psychiatry *Clinical Psychology & Psychotherapy* published online 17 Mar 2011 (DOI: 10.1002/cpp.749)
- Connolly, M.B. & Strupp, H.H. (1996). Cluster Analysis of Patient Reported Psychotherapy Outcomes. *Psychotherapy Research*, 6, 30-42.
- Haggerty, G., Blake, M., Naraine, M., Siefert, C. & Blais, M. (2010). Construct validity of the Schwartz Outcome scale-10: Comparisons to interpersonal distress, adult attachment, alexithymia, the five factor model, romantic relationship length and ratings of childhood memories. *Clinical Psychology and Psychotherapy*, 17(1) 44-50.
- Hilsenroth, M., Ackerman, S., & Blagys, M. (2001). Evaluating the Phase Model of Change During Short-Term Psychodynamic Psychotherapy. *Psychotherapy Research*, 11, 29-47.
- Lambert, M., Hansen, N., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G., & Reinsinger, C. (1996). Administration and Scoring Manual for the Outcome Questionnaire (OQ-45.2). Stevenson, MD. American Professional Credentialing Services. Owens, J., and Imel, Z. (2010). Rating scales in psychotherapy practice. In L. Baer & M. Blais (Eds.), *Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health*. New York, NY: Humana Press, pages 257-270.
- Young, J. L., Waehler, C. A., Laux, J. M., McDaniel, P. S., & Hilsenroth, M. J. (2003). Four studies extending the utility of the Schwartz Outcome Scale (SOS-10). *Journal of Personality Assessment*, 80, 130-138.

Part Three: Performance-Based Tools

The Rorschach Inkblot Method (RIM)

RIM involves the serial presentation of a set of 10 inkblot cards composed of ambiguous and evocative stimuli. Participants are seated side by side with the examiner and are instructed to describe what they see in each card (*What might this be?*). Each response, freely given by the respondent, is scored, after appropriate inquiry, according to the features selected by the subject as relevant. These main features are related to perceptual organization and integration (location, developmental quality and integrative process), characteristics that are actually contained in the card (e.g., form, color, and shading) or are subjectively added to it (e.g., perception of movement). Evaluations are made of the perceptual adequacy of the responses given to the stimuli (form quality), contents of the representation (animal, human figure, anatomical perception, sexual aspects, etc.), and ideation (thought processes in giving the response). All aspects of the response are considered to be projective in that even a popular response to a specific card is given by the participant in his or her own unique or even idiosyncratic way. Reliable scoring criteria are provided for coding each separate aspect.

Although Hermann Rorschach (1921/1942) initially developed his inkblots to assess personality dynamics and refine psychiatric diagnoses, he conceptualized these stimuli primarily as a measure of perceptual style and cognitive organization—that is, as a way to assess people's characteristic ways of perceiving and processing information.

The range of constructs measured by the RIM expanded following Rorschach's death in 1922. Beck (1937) adhered closely to Rorschach's initial emphasis on perceptual and cognitive processes and added various perceptual and information processing scores to the measure, while Klopfer (1937) emphasizing the idiographic and psychodynamic aspects of inkblot responses, developing rules for scoring and interpreting thematic content (e.g., dependency, aggression, preoccupation with control). Rapaport, Gill, and Schafer (1945-1946, 1968) formulated a systematic psychoanalytic approach to the test and

developed a method of analyzing test verbalizations for indications of thought disorder. In this same period, Frank (1939) offered his now-famous *projective hypothesis*, wherein he speculated that unstructured procedures like the RIM compel the respondent to reveal hidden wishes, needs, fears, and motives by projecting these private concerns onto ambiguous stimuli (hence the term “projective test”). Given its early history, it is not surprising that over the years the RIM has come to be seen in multiple ways—as a perceptual task, a problem-solving task, an index of associational patterns, an interpersonal task (because responses must be given to an examiner), and a measure of personality dynamics.

Multiple RIM scoring and interpretation systems evolved during subsequent decades. By the late 1960s, the considerable variability in the validity and clinical utility of scores derived from different RIM scoring methods prompted Exner (1969, 2003) to review the literature in this area, identify those RIM variables with the strongest empirical support, and combine these variables into a single overarching RIM scoring and interpretation system that he termed the *Comprehensive System* (CS). For the next several decades the CS became the most widely used, and studied, RIM scoring and interpretation system, having a far-reaching impact on the assessment of personality and psychopathology and becoming a standard part of psychological test batteries (Bornstein, 2010).

Although many clinicians and researchers continue to use the CS, in 2011 an alternative RIM scoring and interpretation system was created to correct some of the criticisms that had been leveled at the CS (e.g., Wood, Nezworski, Lilienfeld, & Garb, 2003) and to strengthen the psychometric underpinnings of the Rorschach. The *Rorschach Performance Assessment System* (R-PAS; Meyer, Viglione, Mihura, Erard, & Erdberg, 2011) was based in part of the results of a comprehensive meta-analysis of CS variables that identified those scores with the greatest predictive power (Mihura, Meyer, Dumitrascu, & Bombel, 2013). Although the CS and R-PAS represent the two most influential omnibus scoring and interpretation systems for the RIM, a number of other well-validated, clinically useful methods exist that allow for the derivation of narrower RIM indices (e.g., thought disorder, defense style, primary process thinking, object relations; Bornstein &

Masling, 2005; Huprich, 2006).

Rorschach variables have sometimes been grouped into two broad categories—thematic (content) variables (e.g., the proportion of oral dependent imagery in a protocol) and structural (perceptual) variables (e.g., the degree to which the respondent emphasizes common versus unusual details when interpreting the inkblots). However, distinctions between these two categories are less sharp than once thought, and clinical and forensic predictions typically draw on both types of scores (Exner & Erdberg, 2005; Hilsenroth & Stricker, 2004). In both the CS and R-PAS, these variables are summarized in a standardized summary sheet often referred to as a *structural summary*; most RIM interpretations are based on the analysis of combinations of scores from a CS or R-PAS structural summary sheet.

Rorschach scoring and interpretation is complex, requiring considerable formal training and experience. In addition to individual RIM scores, an array of ratios, percentages, and other derivations are used in RIM interpretation. The particular outcome variables that are of greatest interest in a given assessment situation are determined by characteristics of the patient, the purpose of the assessment, and the referral question.

Many techniques in psychodynamic therapy are directed toward increasing introspective access and making unconscious material conscious; an analogous process occurs with the RIM. Because the RIM requires respondents to engage in an unfamiliar task during the testing session, with little guidance and minimal feedback regarding the purpose of the test or the types of outputs produced, the method has long been considered a useful means of assessing aspects of the patient's personality and functioning that might not be accessible to conscious awareness and deliberate self-report.

With this in mind, optimal use of the RIM in psychological assessment requires that RIM scores be interpreted in the context of the patient's performance on tests from other modalities. Exploration of divergences between information that is accessible via introspection and information that shapes behavior unconsciously, implicitly, or reflexively can yield important information regarding underlying personality structure, coping, and defense. For example, research contrasting self-report and RIM indices of interpersonal

dependency have found that patients with features of dependent personality disorder score high on both types of tests whereas those with features of histrionic personality disorder obtain high RIM dependency scores, but low scores on self-report dependency scales.

Strengths. The RIM can be a useful aid in the assessment of M and MA Axis capacities. When administered and scored properly, the RIM yields information on a broad array of psychological domains relevant to personality functioning, unconscious conflicts, and treatment planning. It is useful for assessing children, adolescents, adults, and older adults. It is also sensitive to change in psychotherapy. Substantial clinical and meta-analytical evidence exists (Mihura et al., 2013; Weiner, 2003) supporting the use of RIM variables to assess M Axis functions including:

- cognitive and affective process: organizational processing of stimuli (Z scores), reality testing (XA% and WDA%), ideational process (WSum6), psychotic features (PTI), depressive features (DEPI, MOR), mentalizing activity (M);

- identity and relationships: dependency traits (ROD), object representations (GHR:PHR), interpersonal relationships (CDI, SumH, COP, AG), intimacy needs (SumT);

- defense and coping: stable and situational stress control and tolerance (D scores), adaptation resources (EA), dysphoric feelings (SumShading), some defensive functioning (Lambda index and hypochondriacal concerns), impulse and emotional modulation (WSumC, color ratio, Afr), ego functions (EII-2);

- self-awareness and self-direction: narcissistic-like sense of entitlement (Reflection responses), mature vs. immature identifications (H ratio).

Evidence indicates that the RIM can be of tremendous value in rendering predictions within the clinical setting (Meyer & Handler, 1997), especially for those areas that center on psychological processes largely inaccessible to verbal report (e.g., reality testing, stress tolerance, impulse control). In such situations, the RIM appears to be superior to patient self-report measures, which tap self-attributed traits, motives, and feeling states in lieu of the underlying or implicit processes that impel and direct these responses. Given these

patterns, it is not surprising that RIM scores have been shown to add incremental validity (unique predictive value) in diagnostic and therapeutic settings (Meyer, 2000; Perry, Minassian, Cadenhead, & Braff, 2003; Bram & Peebles, 2014). Moreover, the psychometric properties of scores derived from the RIM are quite strong (Bornstein, 2012): Meta-analytic procedures to estimate retest reliability for a broad array of CS and non-CS RIM scores revealed that both short- and long term retest reliability for RIM scores were good (Grønnerød, 2003, 2006). Other studies have documented the convergent and discriminant validity of RIM scores (Mihura et al., 2013).

Limitations. To be competent in RIM administration, scoring, and interpretation requires considerable training and experience – at least 100 hours of supervised training. Moreover, the RIM is usually lengthy to administer; it is labor-intensive on the part of the patient as well as the examiner. The RIM may not be suitable for very low-functioning patients or those who are dysregulated at the time of testing. For these reasons, as well as because of the complexity of test scoring and interpretation and the method’s strong historical links to psychodynamic theory, the RIM has been less widely used in recent years, and surveys demonstrate that it is less widely taught in graduate training programs than in the past.

An additional limitation of the method concerns RIM interpretation and the ways in which it has occasionally been misused and mislabeled. Contrary to the assertions of critics (e.g., Wood et al., 2005), the RIM is not a diagnostic tool: RIM scores tap underlying psychological processes that are only indirectly related to PDM and DSM symptoms and diagnoses. RIM scores can be used to refine differential diagnoses because certain superficially similar disorders have contrasting underlying dynamics. However, RIM scores are—at best—an adjunct to more traditional diagnostic screening instruments, and by themselves they are not appropriate for making diagnostic classification decisions.

For more information you can see www.r-pas.org or write to Robert F. Bornstein (bornstein@adelphi.edu)

References

- Archer, R. P., & Krishnamurthy, R. (1993a). Combining the Rorschach and MMPI in the assessment of adolescents. *Journal of Personality Assessment*, 60, 132-140.
- Archer, R. P., & Krishnamurthy, R. (1993b). A review of MMPI and Rorschach interrelationships in adult samples. *Journal of Personality Assessment*, 61, 277-293.
- Beck, S. J. (1937). *Introduction to the Rorschach method: A manual of personality study*. NY: American Orthopsychiatric Association.
- Berant, E., Newborn, M., & Orgler, S. (2008). Convergence of Rorschach scales and self-report indexes of psychological distress: The moderating role of self-disclosure. *Journal of Personality Assessment*, 90, 36-43.
- Bornstein, R. F. (2002). A process dissociation approach to objective-projective test score interrelationships. *Journal of Personality Assessment*, 78, 47-68.
- Bornstein, R. F. (2010). Psychoanalytic theory as a unifying framework for 21st century personality assessment. *Psychoanalytic Psychology*, 27, 133-152.
- Bornstein, R. F. (2012). Rorschach score validation as a model for 21st century personality assessment. *Journal of Personality Assessment*, 94, 26-38.
- Bornstein, R. F., & Masling, J. M. (Eds.) (2005). *Scoring the Rorschach: Seven Validated Systems*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bram, A.D. and Peebles, M.J. (2014). *Psychological Testing that Matters: Creating a Road Map for Effective Treatment*. American Psychological Association, Washington, DC.
- Cogswell, A. (2008). Explicitly rejecting an implicit dichotomy: Integrating two approaches to assessing dependency. *Journal of Personality Assessment*, 90, 26-35.
- Cramer, P. (2000). Defense mechanisms in psychology today: Mechanisms for adaptation. *American Psychologist*, 55, 637-646.
- Erdelyi, M. H. (2006). The unified theory of repression. *Behavioral and Brain Sciences*, 29, 499-511.

- Exner, J. E. (1969). *The Rorschach systems*. New York, NY: Grune & Stratton.
- Exner, J.E. (2003). *The Rorschach: A Comprehensive System*. Vol. 1. Basic Foundations and Principles of Interpretation, 4th ed. NY: Wiley.
- Exner, J. E., & Erdberg, P. (2005). *The Rorschach: A comprehensive system, Volume 2: Advanced interpretation* (3rd edition). Hoboken, NJ: Wiley.
- Frank, L.K. (1939). Projective methods for the study of personality. *Journal of Psychology*, 8, 389-413.
- Hilsenroth, M. J., & Stricker, G. (2004). A consideration of challenges to psychological assessment instruments used in forensic settings: Rorschach as exemplar. *Journal of Personality Assessment*, 83, 141-152.
- Huprich, S. K. (Ed.) (2006). *Rorschach Assessment of the Personality Disorders*. Mahwah, NJ: Erlbaum.
- Klopfer, B. (1937). The present status of the theoretical developments of the Rorschach method. *Rorschach Research Exchange*, 1, 142-147.
- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, 96, 690-702.
- Meyer, G. J. (2000). The incremental validity of the Rorschach Prognostic Rating Scale over the MMPI Ego Strength Scale and IQ. *Journal of Personality Assessment*, 74, 356-370.
- Meyer, G. J., & Handler, L. (1997). The ability of the Rorschach to predict subsequent outcome: A meta-analysis of the Rorschach Prognostic Rating Scale. *Journal of Personality Assessment*, 69, 1-38.
- Meyer, G. J., Viglione, D. J., Mihura, J. L., Erard, R. E., & Erdberg, P. (2011). *Rorschach Performance Assessment System: Administration, coding, interpretation, and technical manual*. Toledo, OH: Rorschach Performance Assessment System, LLC.
- Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the Comprehensive System. *Psychological Bulletin*, 139, 548-605.
- Perry, W., Minassian, A., Cadenhead, K., & Braff, D. (2003). Use of the Ego Impairment

- Index across the schizophrenia spectrum. *Journal of Personality Disorders*, 80, 50-57.
- Rapaport, D., Gill, M. M., & Schafer, R. (1945-1946). *Diagnostic psychological testing*. Chicago, IL: Yearbook Publishers.
- Rapaport, D., Gill, M. M., & Schafer, R. (1968). *Diagnostic psychological testing* (2nd ed.; R. R. Holt, ed.). New York, NY: International Universities Press.
- Rorschach, H. (1942). *Psychodiagnostics: A diagnostic test based on perception* (P. Lemkau & B. Kronenberg, trans.). Berne, Switzerland: Huber. (Original work published 1921)
- Weiner, I. B. (2000). Making Rorschach interpretation as good as it can be. *Journal of Personality Assessment*, 74, 164-174.
- Weiner, I.B. (2003). *Principles of Rorschach interpretation*. 2nd ed. Mahwah, NJ: Lawrence Erlbaum.
- Wood, J. M., Nezworski, M. T., Lilienfeld, S. O., & Garb, H. N. (2003). *What's wrong with the Rorschach?* San Francisco, CA: Jossey-Bass.

The Thematic Apperception Test (TAT)

The TAT was the creation of artist Cristiana Morgan and psychiatrist Henry Murray (Morgan & Murray, 1935). Murray (1943, p. 3) described the TAT as “a method of revealing some of the dominant drives, emotions, sentiments, complexes, and conflicts of personality”. Three contributions of the TAT response - card stimulus, testing environment, and the patient’s inner world – help account for individual differences (Aronow, Weiss, & Reznikoff, 2001). It is typically left up to the examiner to differentiate what aspects of an individual’s responses can be attributed to each of these sources (Stein et al., 2013).

The TAT consists of 31 stimulus cards (1,2, 3BM, 3GF, 4, 5, 6BM, 6GF, 7BM, 7GF, 8BM, 8GF, 9BM, 9GF, 10, 11, 12M, 12F, 12BG, 13MF, 13B, 13G, 14, 15, 16, 17BM, 17GF, 18BM, 18GF, 19, 20). The letters following the number were originally meant to correspond to the gender and age of the person (e.g., young boy and girl, male and female over 14). However, since its development, there is increased flexibility regarding choosing

both the type and number of cards. Some clinicians choose the card based on the gender and age of the patient, and TAT protocols vary as a result. Other clinicians choose cards on the basis of thematic content, regardless of the age. Likewise, protocols can vary in response to the clinical question about the patient, whereas some clinicians maintain a standard set of TAT cards across patients. The lack of a specific set of rules regarding the construction of a TAT protocol creates a challenge for researchers who seek to validate scores derived from the TAT and to document its clinical utility empirically.

The cards consist of single to multiple figures in various interpersonal situations, all of which evoke a mixture of self and interpersonal themes and are “differently arousing” (Ehrenreich, 1990). Individuals are asked to make up a story around the picture, with a beginning, middle, and ending. Also, individuals are asked to provide what the character(s) may be thinking and feeling.

There are two methods of TAT interpretation: nomothetic (using an empirically validated scoring system and comparing responses to normative data) and idiographic (examining individual responses and discerning meaning about what these responses express about underlying personality structure). Clinicians are encouraged to use a combination of both approaches when interpreting TAT protocols (Aronow et al, 2001).

The two most common measures used to rate TAT narratives are the Social Cognition and Object Relations Scale (SCORS; Westen, 1995, see above) and the Defense Mechanism Manual (DMM; Cramer, 1991, see above). Historical accounts of the TAT by major contributors (e.g., Leopold Bellak, Robert Holt, David McClelland, Edward Schneiderman, Saul Rosenzweig, Morris I. Stein) to its literature can be found in Gieser and Stein (1999). Jenkins (2008) created a comprehensive handbook of TAT scales that have been used in the literature.

Clinicians focus more on the thematic content (see, e.g., Schafer, 1958), rather than a specific scoring system, to code narrative data; they usually focus on dominant themes as well as on how responses deviate from the stimulus pull of the card (Stein et al., 2013). The more idiosyncratic the response, the more TAT narratives may reflect aspects of psychopathology or personality. Also, how a patient approaches the TAT (i.e., responding

to ambiguous, emotionally arousing interpersonal situations) and the patient's behavior during testing can be helpful in understanding the patient's internal world (Aronow et al., 2001).

Strengths. The TAT can be a useful measure in assessing implicit processes that are not readily observed or explicitly stated, such as thought process, interpersonal and object relational themes, dominant affects or emotions, defensive functioning, and psychological conflicts (intra and interpersonal). Moreover, many of the M and MA Axis levels can be evaluated with the TAT; however some of these levels are easier to assess than others. For example, SCORS-G ratings of TAT narratives can provide information regarding most mental functions that are captured in M and MA Axes (with the exception of defensive functioning). M and MA Axes level "Defensive patterns and capacities" can be formally assessed with the DMM (see above) or freely by a clinician familiar with defensive functioning. Evaluating M Axis' "capacity for affective experience, expression, and communication" is relatively easy to accomplish, and does not require a formal coding system.

The TAT permits much flexibility regarding card set, number of cards per protocol, and type of interpretative approach. Clinicians can choose from a variety of scales to code TAT narratives in accordance with the clinical goals of the testing. Identification of themes permits ready comparison with other test data. There have been numerous studies demonstrating its clinical utility and value (e.g., Fowler Ackerman, Speanburg, Bailey, & Blagys, 2004), including after psychodynamic treatment. TAT scores and content are sensitive to treatment changes; for example, patient's complexity of representations of people (as evidenced in TAT narratives) increased (medium effect size) over the course of treatment.

Research had has demonstrated the utility of the TAT in assessing defense mechanisms across diagnostic groups (Cramer & Kelly, 2004), between genders (Cramer, 2002), in longitudinal contexts (Cramer, 2012), and in the context of therapeutic change (Cramer & Blatt, 1990). In more recent years, the TAT has also been used as a tool to assess social

processing deficits caused by brain abnormalities (Paul, Schieffer, & Brown, 2004).

Limitations. Administration of the TAT takes much time and can often be complex, especially with participants who have difficulty in constructing and telling stories around its pictures. The lack of a standard set of TAT cards to administer limits generalizability in clinical research, as does the lack of a widely accepted TAT scoring system, such as the Comprehensive System for the Rorschach (Exner, 1995). The vast majority of clinicians using and interpreting the TAT rely on their clinical judgment, with all of its attendant strengths and established limitations. When comprehensive scoring systems are used to interpret the TAT, they are time consuming and require considerable training. There has been less research differentiating stimulus pull and individual differences, which have notable clinical implications (Stein et al., 2013). There have also been questions regarding cultural sensitivity and the dated content of cards. For all of these reasons, as well as because projective techniques are so strongly linked to psychodynamic theory, surveys show that psychologists are using the TAT (and other performance-based measures) less frequently than in the past, and graduate training programs are teaching it less frequently too.

For more information, you can write to Michelle Stein (steinmischelleb@yahoo.com).

References

- Ackerman, S., Hilsenroth, M., Clemence, A., Weatherill, R., & Fowler, C. (2000). The Effects of Social Cognition and Object Representation on Psychotherapy Continuation. *Bulletin of the Menninger Clinic, 64*, 386-408.
- Aronow, E., Weiss, K. A., & Reznikoff, M. (2001). A practical guide to the Thematic Apperception Test: The TAT in clinical practice. New York, NY: Routledge.
- Cramer, P. (2012). Psychological maturity and change in adult defense mechanisms. *Journal of Research in Personality, 46*, 306-316.
- Cramer, P. & Kelly, F.D. (2004). Defense mechanisms in adolescent conduct disorder and

- adjustment reaction. *Journal of Nervous and Mental Disease*, 192(2), 139-145.
- Cramer, P. (2002) The study of defense mechanisms: Gender implications, In Bornstein, R. F., & Masling, J. M. E. (Ed.) *The psychodynamics of gender and gender role* (pp. 81-127). Washington, DC: American Psychological Association.
- Cramer, P. (1991). The development of defense mechanisms: Theory, research and assessment. New York, NY: Springer-Verlag.
- Cramer, P., & Blatt, S.J. (1990). Use of the TAT to measure change in defense mechanisms following intensive psychotherapy. *Journal of Personality Assessment*, 54 (1 & 2), 236-251.
- Ehrenreich, J. H. (1990). Quantitative studies of responses elicited by selected TAT cards. *Psychological Reports*, 67, 15–18.
- Fowler, J.C., & DeFife, J.A. (2012). Quality of Object Representations Related to Service Utilization in a Long-Term Residential Treatment Center. *Psychotherapy*, 49(3), 418-422.
- Fowler, C., Ackerman, S., Speanburg, S., Bailey, A., Blagys, M. & Conklin, A.C. (2004). Personality and symptom change in treatment-refractory inpatients: Evaluation of the phase model of change using the phase model of change using Rorschach, TAT and DSM-IV Axis V. *Journal of Personality Assessment*, 83, 306-322.
- Gieser, L, & Stein, M. I. (Eds.), (1999). *Evocative images: The Thematic Apperception Test and the art of projection*. Washington, DC: American Psychological Association.
- Jenkins, S. R. (2008). A handbook of clinical scoring systems for the thematic apperceptive techniques. New York, NY: Taylor & Francis.
- Morgan, C. D., & Murray, H. A. (1935). A method of investigating fantasies: The Thematic Apperception Test. *Archives of Neurology and Psychiatry*, 34, 289–306.
- Paul, L.K., Schieffer, B., and Brown, W.S. (2004) Social Processing Deficits in Agenesis of the Corpus Callosum: Narratives from the Thematic Apperception Test. *Archives of Clinical Neuropsychology*. 19, 215-225.
- Schafer, R. (1958). How was this story told. *Journal of Projective Techniques*, 22, 181-210.

Stein MB, Slavin-Mulford J, Siefert CJ, Sinclair SJ, Smith M, Chung WJ, Liebman R, & Blais MA (2015) External validity of SCORS-G ratings of Thematic Apperception Test narratives in a sample of outpatients and inpatients. *Rorschachiana*, 36 (1)

Stein, M.B., Slavin-Mulford, J., Siefert, C.J., Sinclair, S.J., Renna, M., Malone, J., Bello, I., & Blais, M.A. (2013). SCORS-G Stimulus Characteristics of Select Thematic Apperception Test Cards. *Journal of Personality Assessment*, DOI: 10.1080/00223891.2013.823440

Part Four: Instruments evaluating the treatment

The following instruments have utility for assessing treatments, for determining how the patient is functioning with a particular therapist, and also for evaluating the patient's psychological functioning and health directly in this particular setting. Furthermore, changes in the way the patient and therapist are functioning over the course of treatment also are likely to reflect changes in psychological health and to constitute a measure of interpersonal functioning. Of course, the functioning of the therapeutic dyad also reflects the skill and relational capacity of the therapist in general and with that individual patient in particular. Nevertheless, many studies have provided evidence that patient characteristics play an important role in variation in the therapy relationship and in outcome (Norcross, 2011).

The Working Alliance Inventory (WAI)

The WAI (Horvath, 1981, 1982; Horvath & Greenberg, 1986, 1989) is a widely used instrument assessing the therapeutic alliance on the basis of Bordin's (1979) pantheoretical model of the therapeutic alliance as involving a mutual agreement between client and clinician on the goals of treatment and the tasks necessary to achieve them, together with the establishment of a relational bond that maintains the collaboration between participants (patient and therapist) in the therapeutic work. It consists of 36 items rated on a 7-point, Likert-type scale ranging from 1 (never) to 7 (always). The inventory generates three 12-item subscales for (a) the goal, (b) the task, and (c) the bond.

Horvath and his colleagues developed three versions of the WAI: a client form (WAI-P); a therapist form (WAI-T); and an observer form (WAI-O). Although the WAI-P and WAI-T are administered at the end of a therapy session, the WAI-O is used for rating psychotherapy session transcripts. Several briefer forms have been developed as well. A short form, the WAI-S (Tracey & Kokotovic, 1989) comprises 12 items and is available in two versions (client and therapist). A short version of the WAI was also developed and

validated by Hatcher and Gillaspay (2006) and is widely used.

Internal consistency of the three subscales of the WAI-P and WAI-T are high, as are internal consistency and interrater reliability values of the WAI-O's subscales (Horvath & Greenberg, 1986, 1989; Tichenor & Hill, 1989). In general, the three dimensions (bond, task, goal) are strongly correlated (Horvath & Greenberg, 1989; Plotnicov, 1990; Safran & Wallner, 1991). Despite the high correlations, they are distinct but overlapping dimensions (Tracey & Kokotovic, 1989).

Strengths. The WAI is a valid and reliable measure of the working alliance that can be administered easily and completed rapidly, whether by a client or a therapist or a nonparticipant observer. It has been used by researchers and practitioners of all theoretical orientations interested in exploring how the alliance interacts with, and affects treatment process and outcome. Many studies have shown a strong relationship between working alliance and benefit from treatment (e.g. Norcross, 2011); indeed, the alliance as viewed by the patient and as measured at an early session is a robust predictor of and contributor to treatment outcome in all forms of psychotherapy evaluated.

Limitations. The WAI (all the versions) measures the quality of the therapeutic alliance at a macro level and may not prove useful for assessing in-session therapeutic alliance fluctuations (e.g., alliance ruptures and resolutions) through a microanalytic investigation of the way patient and therapist co-construct their relationship. WAI scores, like self-report measures in general, can be easily manipulated or faked by the patient.

For more information, see <http://wai.profhorvath.com/>.

References

Hatcher, R. L. & Gillaspay, J. A. (2006). Development and Validation of a Revised Short Version of the Working Alliance Inventory (WAI-SR). *Psychotherapy Research*, 16, 12-25.

- Horvath, A. O. (1981). An exploratory study of the working alliance: Its measurement and relationship to outcome. Unpublished doctoral dissertation, University of British Columbia, Vancouver, Canada.
- Horvath, A. O. (1982). Working Alliance Inventory (revised). Unpublished manuscript, Simon Fraser University, Burnaby, British Columbia, Canada.
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36, 223–233.
- Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, 38, 139–149.
- Norcross, J C [Ed]. *Psychotherapy relationships that work: Evidence-based responsiveness*. 203-223. New York, NY, US: Oxford University Press; US.)
- Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 207–210.

The Therapist Response Questionnaire (TRQ)

The Therapist Response Questionnaire (TRQ; Zittel Conklin & Westen, 2003; Betan, Heim, Zittel Conklin, & Westen, 2005), previously called the Countertransference Questionnaire, is a 79-item clinician report designed to assess the clinician’s emotional reactions to patients in psychotherapy.

The items measure a wide spectrum of thoughts, feelings, and behaviors expressed by therapists toward their patients, ranging from relatively specific feelings (e.g., “I feel bored in sessions with him/her”) to more complex constructs (e.g., “More than with most patients, I feel like I’ve been pulled into things that I didn’t realize until after the session was over”). They were derived by reviewing the clinical, theoretical, and empirical literature on countertransference and related variables, and are written in straightforward jargon-free terms near to clinical experience, so that clinicians of any theoretical orientation can easily understand them. The therapists are asked to assess each item on a 5-point Likert scale ranging from 1 (not true) to 5 (very true).

The TRQ's factor structure comprises eight emotional dimensions that are clinically sensitive and conceptually coherent: overwhelmed/disorganized, helpless/inadequate, positive, special/overinvolved, sexualized, disengaged, parental/protective, criticized/mistreated.

These scales showed excellent internal consistency and good criterion validity. Several studies (see, for example, Colli et al., 2014; Gazzillo et al., 2015) showed that therapists' emotional response patterns are related in predictable ways to patients' personality pathology across therapeutic approaches, and that clinicians, regardless of therapeutic orientation, can make diagnostic and therapeutic use of their own emotional responses to patients.

The adolescent version of the TRQ (Satir, Thompson-Brenner, Boisseau, & Crisafulli, 2009) consists of 86 items, and its factor structure comprises six emotional response dimensions similar to those of the TRQ for adults: angry/frustrated, warm/competent, aggressive/sexual, failing/incompetent, bored/angry at parents, and overinvested/worried.

The TRQ may be a useful aid for the assessment of Axes P and PA.

Strengths: The TRQ is a valid and reliable instrument for the assessment of clinicians' emotional responses. It is easy to score and can be used for both clinical and research purposes.

Limitations: The main limitation of the TRQ is the possible influence of social desirability bias or implicit defensive processes affecting all self-report measures. Empirical investigations of clinicians' emotional responses could benefit from the research designs including other methods of measurement and different perspectives (e.g., external observer or supervisor).

References

Betan, E., Heim, A. K., Zittel Conklin, C., & Westen, D. (2005). Countertransference phenomena and personality pathology in clinical practice: An empirical investigation. *The American Journal of Psychiatry*, 5, pp. 890–898.

- Colli A., Tanzilli A., Dimaggio G., & Lingardi V. (2014). Patient personality and therapist response: An empirical investigation. *The American Journal of Psychiatry*, 171, pp. 102–108.
- Gazzillo, F., Lingardi, V., Del Corno, F., Genova, F., Bornstein, R. F., Gordon, R. M., & McWilliams, N. (2015). Clinicians' emotional responses and Psychodynamic Diagnostic Manual adult personality disorders: A clinically relevant empirical investigation. *Psychotherapy*, 52(2), pp. 238–246.
- Lingardi, V., Tanzilli, A., & Colli, A. (2015). Does the severity of psychopathological symptoms mediate the relationship between patient personality and therapist response?. *Psychotherapy*, 52(2), pp. 228–237.
- Satir, D. A., Thompson-Brenner, H., Boisseau, C. L., & Crisafulli, M. A. (2009). Countertransference reactions to adolescents with eating disorders: Relationships to clinician and patient factors. *International Journal of Eating Disorders*, 42(6), pp. 511–521.

The Psychotherapy Relationship Questionnaire (PRQ)

The PRQ (Westen, 2000; Bradley, Heim, & Westen, 2005) is a 90-item clinician-report questionnaire that assesses patients' interpersonal patterns in the relationship with his or her therapist.

The PRQ items measure a wide range of thoughts, feelings, motives, conflicts and behaviors expressed by patients toward their therapist. They were derived by a review of the clinical, theoretical, and empirical literature on transference, therapeutic or working alliance, and related constructs and are written in everyday language, without jargon, so that the instrument can be used by clinicians of any theoretical approach. The therapists have to assess each item on a 5-point Likert scale, ranging from 1 (not true) to 5 (very true).

The PRQ's factor structure comprises five transference dimensions that are clinically and theoretically coherent: angry/entitled, anxious/preoccupied, secure/engaged, avoidant/counterdependent, and sexualized. The PRQ's scales showed excellent internal consistency and good criterion validity that is systematically related to personality pathology. Additionally, four of the five relational patterns of PRQ are similar to the adult attachment styles identified using the Adult Attachment Interview (AAI; Main et al , 1985).

PRQ can be a useful aid in the assessment of P, PA, M and MA Axes (in particular,

relational abilities).

Strengths: The PRQ is a psychometrically robust instrument for the assessment of patients' relational patterns emerging in the therapeutic relationship; it provides a window on patients' personality and relational capacities in general that is based upon direct observation. It is easy to administer and relatively brief.

Limitations: The PRQ is subject to any inherent biases (countertransference) of the clinician. Therefore any findings would be strengthened by ratings by an external observer as well.

References:

Bradley, R., Heim, A.K., & Westen, D. (2005). Transference patterns in the psychotherapy of personality disorders: empirical investigation. *The British Journal of Psychiatry*, 186, pp. 342-349.

The Assimilation of Problematic Experiences Sequence (APES)

The APES summarizes a developmental progression that patients experience as they overcome problems in successful psychotherapy, according to the assimilation model (Stiles et al., 1991, 2005, 2011). Specifically, it describes the assimilation of problematic experiences, such as traumas or dysfunctional relationships, into the patient's normal schemas. After the experiences are assimilated, they are no longer problematic but become resources, available as called upon to meet life's challenges and opportunities. For example, uncontrolled verbal outbursts may become assimilated as a capacity for assertiveness (Stiles, 1999).

The APES is construed as a continuum anchored by eight stages, numbered 0 to 7, that characterize the changing relation of an initially problematic experience to the rest of the self: (0) warded off/dissociated, (1) unwanted thoughts/active avoidance, (2) vague awareness/emergence, (3) problem statement/clarification, (4) understanding/insight, (5) application/working through, (6) resourcefulness/problem solution, and (7) integration/mastery. In successful psychotherapy, problematic experiences move through

some segment of this sequence.

Stages of the APES can be rated from any clinical material, including audio or video recordings, transcripts, session summaries, or patient-produced materials. Rater selection and training procedures have varied a good deal across studies, and there exist several versions of this tool, qualitative and quantitative (e.g., Caro Gablada, 2006, 2009; Gray & Stiles, 2011; Osatuke & Stiles, 2011; Tikkanen, Stiles, & Leiman, 2011, 2013;), for different kinds of pathologies, from psychological to neurological problems and mental retardation, and for different treatment modalities and settings, from individual to couple, family and group therapy, from psychotherapy to pharmacotherapy and supervision (e.g., Cheston, 2013; Osatuke & Stiles, 2012; Osatuke, Reid, Stiles, Zisook, & Mohamed, 2011).

The relation of the APES to symptom intensity measures has been studied. Theoretically, distress and symptom intensity vary systematically across APES stages, but the relation is not linear (Stiles, Osatuke, Glick, & Mackay, 2004). Patients who enter therapy with problematic experiences that are warded off or avoided (APES stages 0 or 1) are likely to feel worse before they feel better, as these experiences emerge and are acknowledged and confronted. The most intense and sustained emotional pain is expected at APES stage 2 (vague awareness/emergence). Progress through APES stages 2-6 is characterized by decreasing distress (stages 2-4) and then increasing positive affect (stages 4-6), with the most rapid change across stages 3-5, as the patient identifies, labels, and formulates the problem, works toward an understanding, and explores applications of the understanding in everyday life. The consolidation at APES stages 6-7 is expected to involve integrating and normalizing and may involve decreases in the euphoric feeling associated with solving a problem.

As recognized not only in psychodynamic therapies but across many theoretical approaches, insight and understanding are intimately linked with psychotherapy outcome, assessed as symptom intensity reduction (Castonguay & Hill, 2007). Understanding of conventional outcome criteria via the assimilation model thus suggests that APES stage 4 (understanding/insight) represents the point of greatest reduction in distress and symptom intensity. Consistent with this analysis, APES stage 4 appears pivotal in distinguishing

conventionally-assessed good-outcome cases from poor-outcome cases (Detert et al., 2006).

Strengths. The APES can be useful for assessing Axes I, II, and III dimensions such as mentalization, self-observation, and resiliency. It is applicable to a broad range of patient populations and treatment settings. The APES goes beyond conventional symptom-intensity-based outcome measures to assess the internal dynamics of psychological change, sometimes described as structural change, as used by the Operationalized Psychodynamic Diagnosis (OPD) group (see above in Section 2; and Grande, Rudolf, Oberbracht, & Jakobsen, 2001; Grande, Rudolf, Oberbracht, & Pauli-Magnus, 2003;). It links to an evolving description of the process of psychological change and can be applied to a broad range of clinical material. There are many overlaps of potential interest between the APES and other measures of personality described in this chapter, such as the SCORS, the ORI, RF, and the APS measures of patient functioning in the therapeutic hour.

Limitations. Considerable clinician time is required to rate the clinical material according to the APES, which has rarely been used for routine clinical use. Indeed, the APES is not a psychological test or an instrument per se. Most of its research comprises smaller, qualitative studies. The APES is an evolving tool, and there are many versions of procedures for applying it. Although all aim at describing the same developmental sequence, there is not a single standard procedure.

For more information, you write to William B. Stiles (stileswb@miamioh.edu) or see <http://www.users.miamioh.edu/stileswb/>.

References

Barkham, M., Rees, A., Stiles, W. B., Hardy, G. E., & Shapiro, D. A. (2002). Dose-effect relations for psychotherapy of mild depression: A quasi-experimental comparison of

- effects of 2, 8, and 16 sessions. *Psychotherapy Research*, 12, 263-274.
- Castonguay, L. G., & Hill, C. E. (Eds.) (2007). *Insight in psychotherapy*. Washington, DC: American Psychological Association.
- Detert, N. B., Llewelyn, S. P., Hardy, G. E., Barkham, M., & Stiles, W. B. (2006). Assimilation in good- and poor-outcome cases of very brief psychotherapy for mild depression: An initial comparison. *Psychotherapy Research*, 16, 393-407.
- Grande, T., Rudolf, G., Oberbracht, C., & Jakobsen, T. (2001). Therapeutic changes beyond the symptoms: Effects of inpatient treatment in the view of the Heidelberg Structural Change Scale. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 47, 213–233.
- Grande, T., Rudolf, G., Oberbracht, C., & Pauli-Magnus, C. (2003). Progressive changes in patients' lives after psychotherapy: Which treatment effects support them? *Psychotherapy Research*, 13, 43-58.
- Honos-Webb, L., Stiles, W. B., & Greenberg, L. S. (2003). A method of rating assimilation in psychotherapy based on markers of change. *Journal of Counseling Psychology*, 50, 189–198.
- Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, 62, 1009-1016.
- Reid, M., & Osatuke, K. (2006) Acknowledging problematic voices: Processes occurring at early stages of conflict assimilation in patients with functional somatic disorder. *Psychology and Psychotherapy: Theory, Research and Practice*, 79, 539-555.
- Rudolf, G., Grande, T., & Oberbracht, C. (2000). The Heidelberg restructuring scale. A model of changes in psychoanalytic therapies and its operationalization on an estimating scale. *Psychotherapeut*, 45, 237–246.
- Stiles, W. B. (1999). *Signs, voices, meaning bridges, and shared experience: How talking helps*. Visiting Scholar Series No. 10 (ISSN 1173-9940). Palmerston North, New Zealand: School of Psychology, Massey University.
- Stiles, W. B. (2002). Assimilation of problematic experiences. In J. C. Norcross (Ed.),

Psychotherapy relationships that work: Therapist contributions and responsiveness to patients (pp. 357-365). New York: Oxford University Press.

Stiles, W. B. (2005). Extending the Assimilation of Problematic Experiences Scale:

Commentary on the special issue. *Counselling Psychology Quarterly*, 18, 85-93.

Stiles, W. B. (2011). Coming to terms. *Psychotherapy Research*, 21, 367-384

Stiles, W. B., Morrison, L. A., Haw, S. K., Harper, H., Shapiro, D. A., & Firth-Cozens, J.

(1991). Longitudinal study of assimilation in exploratory psychotherapy.

Psychotherapy, 28, 195-206.

Stiles, W. B., Osatuke, K., Glick, M. J., & Mackay, H. C. (2004). Encounters between internal voices generate emotion: An elaboration of the assimilation model. In H. H.

Hermans & G. Dimaggio (Eds.), *The dialogical self in psychotherapy* (pp. 91-107).

New York: Brunner-Routledge.

Tikkanen, S., Stiles, W. B., & Leiman, M. (2013). Achieving an empathic stance:

Dialogical sequence analysis of a change episode. *Psychotherapy Research*, 23,

178-189.

Varvin, S. (2003). *Mental survival strategies after extreme traumatisation*. Copenhagen,

Denmark: Multivers APS Publishers.

Watkins, R., Cheston, R., Jones, K., & Gilliard, J. (2006). 'Coming out' with Alzheimer's

disease: Changes in awareness during a psychotherapy group for people with dementia.

Aging & Mental Health, 10, 166-176.

The Psychotherapy Process Q-Set (PQS)

The PQS (Jones, 1985, 2000) consists of 100 items assessed through a Q-sort method (Block, 1961/1978). PQS items cover a wide range of dimensions of the psychotherapy process, including both relational and technical aspects. Moreover, the PQS items describe patient contributions to the psychotherapy process (e.g., Q97: Patient is introspective, readily explores inner thoughts and feelings), therapist contributions (e.g., Q50: Therapist draws attention to feelings regarded by the patient as unacceptable, such as anger, envy, or excitement), and patient-therapist interactions (e.g., Q39: There is a competitive quality to

the relationship).

The PQS provides a description of the psychotherapy process, both of adult and adolescent treatment, suitable for comparison and quantitative analysis (Jones, 2000). After studying the transcripts of a therapy hour, clinical judges proceed to the ordering of the 100 items. These were originally printed separately on cards to permit easy arrangement and rearrangement and are sorted into nine piles on a continuum from least (category 1) to most characteristic (category 9). More recently, this sort is more readily accomplished with a computer program. PQS ratings, as other Q-sort measures, are based on forced choice, so that the number of items scored at each level is fixed (e.g., ranging from five items required at the extremes to 18 in the middle category) and conforms approximately to a normal distribution.

The reliability of the instrument is strong, with alpha coefficients ranging from .83 to .92 with evaluators of different theoretical orientations (Jones, Hall, & Parke, 1991; Jones & Pulos, 1993). PQS results were reliable in the evaluation of session transcripts from treatments of different orientations and were successful in differentiating between types of therapies.

An early fruit of this labor was the demonstration that in samples of both psychodynamic and cognitive-behavioral treatments, better results were associated with the psychodynamic aspects of treatment than the cognitive behavioral aspects (Ablon & Jones, 1998).

Strengths. The portrayal of the patient-therapist relationship emerging from the PQS assesses the patient's capacities for relatedness, helping the assessment of P and PA Axes and, more particularly, of the capacity of relationship and intimacy of M and MA Axes. The PQS enables a qualitative and quantitative assessment of psychotherapy process and leads to both a global and a detailed picture of a session and a period of therapy.

Like other Q-set procedures, the PQS prevents scores from being biased upwards or downwards by temperament and is based on the assessment of several different dimensions of patient and therapist contributions to the therapeutic process.

The PQS can be useful in capturing the richness and complexity of psychotherapy, as well as the particular nature of the patient and therapist relationship. A careful recognition and understanding of these repetitive interactions is useful in clinical practice because these patterns can be linked to positive or negative therapy outcome (Josephs et al. 2014).

Limitations. The PQS requires recorded sessions, from which transcripts are made, in order to evaluate the treatment systematically. The PQS assessment is time consuming and needs training to be reliable. In general, 90 minutes are needed to assess a 50-minute session, with reading and listening to the session included. For a reliable assessment, the average length of the PQS training is 25 hours. Also, some of the profiles of types of treatment are now regarded as reflecting certain outdated psychoanalytic concepts, such as neutrality, abstinence, and avoiding expression of one's own subjectivity. As a consequence, findings about the efficacy of different profiles of practice using the PQS can be misleading.

For more information, write to Stuart Ablon (sablon@partners.org)

References

- Ablon, S. & Jones, E. E. (1998). How Expert Clinicians' Prototypes of an Ideal Treatment Correlate with Outcome in Psychodynamic and Cognitive-behavioral Therapy. *Psychother. Research*, 8, 71-83.
- Ablon, S. & Jones, E. E. (2002). Validity of Controlled Clinical Trials of Psychotherapy: Findings from the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry*, 159, 775–783.
- Ablon, S. J., & Jones, E. E. (2005). On Analytic Process. *Journal of the American Psychoanalytic Association* 53, 541-568
- Ablon, J. S., Levy, R.A., & Katzenstein, T. (2006). Beyond Brand Names of Psychotherapy: Identifying Empirically Supported Change Processes. *Psychotherapy: Theory, Research, Practice, Training*.43, 216-231.

- Block, J., (1961). The Q-sort method in personality assessment and psychiatric research. , (pp. 3-26). Springfield, IL, US: Charles C. Thomas Publisher.
- Jones, E.E. (2000). *Therapeutic Action: a Guide to Psychoanalytic Therapy*. Northvale, NJ: Jason Aronson.
- Jones, E. E., & Windholz, M. (1990) The Psychoanalytic Case Study. Toward a Method for Systematic Inquiry. *Journal of the American Psychoanalytic Association*: 39, 985-1016.
- Josephs, L., Sanders, A., & Gorman, B.S. (2014). Therapeutic interaction with an older personality disordered patient. *Psychodynamic Psychiatry*:42, 151-172.
- Levy, R., Ablon, J., Thomä, H., Kächele, H., Ackerman, J., Erhardt, I., & Seybert, C. (2012). A specimen session of psychoanalytic therapy under the lens of the Psychotherapy Process Q-set. In R. Levy, S. Ablon & H. Kächele (Eds.), *Psychodynamic Psychotherapy Research* (pp. 509-528). New York: Humana Press.

The Analytic Process Scales (APS)

The APS (Waldron et al., 2004) was developed by a research group of experienced psychoanalysts beginning in 1985 (APS Research Group). There are 32 scales (scored from 0–4) aimed at the assessment of different dimensions of the therapeutic process on the basis of audio-recorded and transcribed sessions of psychotherapy. The patient's contribution to the therapy is assessed via 14 scales, and the therapist's contribution via 18 scales. Definitions and examples of each of these scales are assembled into an 81-page coding manual (Scharf et al., 2010). The anchoring made possible by the APS coding manual facilitates reliable measurements and has also been useful to students of psychotherapy because it combines definitions of core psychoanalytic constructs with clinical examples.

The patient scales assess the degree to which the patient is able to convey his or her experiences, self-reflect on them, and convey his or her feelings, both in regard to the therapist and therapeutic relationship and in regard to other relationships; the extensiveness of his or her communications about romantic and sexual themes, assertiveness,

aggressiveness, hostility, self-esteem, and developmental experiences; the degree to which he or she able to respond to the therapist in a useful manner; and the overall productivity of his or her communications. That is, the focus of assessment is the degree to which the patient's communications show a deepening in self-awareness, contact with own feelings, and cooperation with the therapist.

The therapist scales assess different kinds of intervention (encouraging elaboration, clarification, interpretation, and support), different targets of these interventions (defenses or resistances, transference, and conflicts), and different domains addressed (romantic and sexual life, self-esteem, and development). Therapist scales enable also the assessment of the degree to which the therapist communications are shaped by his or her feelings and are confrontational, amicable, or hostile. Finally, a scale assesses the overall goodness of therapist interventions—that is, their aptness in kind, content, language, and timing.

The APS variables were originally applied to segments of each session, but the procedure has been expanded to apply to each session in its entirety, so that a much larger number of sessions could be studied in a given time.

Reliability has been demonstrated in rating of segments (Waldron et al., 2004b; Lingiardi et al., 2010) and whole sessions (Waldron et al., 2013). In the whole-session assessment procedure, the scorer estimates both an average and the highest level reached in that session on any given variable. The average intraclass correlation (ICC) of the APS patient scales applied by three trained independent raters to 120 sessions was .76, ranging from .67 to .83; the average ICC for APS therapist scales was .77, ranging from .63 to .85 (Gazzillo et al., 2014)

A replicated finding has been that the rated overall quality of therapist communications is the most powerful element in predicting benefit, irrespective of the particular nature of the communication (Waldron et al., 2004a, Waldron & Helm, 2004, Lingiardi et al., 2010, Gazzillo et al., 2014).

Strengths. The APS provides measures of the patient's ability to form a productive relationship with the therapist and to be thoughtful about his or her mental life, in this way

assisting in the assessment of the patient's capacities in the M Axes such as the capacity to self-reflect and to express and communicate feelings. The APS assesses core dimensions of patients' and therapists' contribution to the dynamic processes of psychotherapy and assesses as well the *quality* of the therapist's contribution, a rare feature in instruments applied to psychotherapies. Assessing the segments of each session permits exploration of the moment-to-moment impact of therapist's communications on the patient and vice versa. It most directly addresses psychodynamic activities of the therapist, such as encouraging elaboration, clarifying, interpreting, and addressing defenses, conflicts, and transference reactions and the impact of the developmental years, so the ratings facilitate exploring connections between therapeutic activity and patient changes. The APS therapist scales can help answer questions about the relative contribution of technique and the relationship to treatment outcomes (Waldron et al., 2013), while the APS patient scales can help assess several M axis capacities, such as self-reflection and the expression of feelings.

Limitations. The segmental assessment of the APS is time consuming (for a 50-minute session, the segmental coding takes about 4 hours) and therefore is a procedure to be reserved for intensive study. The whole-session assessment is less time consuming: 60 minutes to assess a 50 minute session (listening while reading the transcript included), but does not permit the same level of detailed exploration of the relationship between therapist communication and responses from the patient. The measure has not so far been used nearly so widely as is the PQS.

References

- Gazzillo, F., Waldron, S., Genova, F., Angeloni, F., Ristucci, C., & Lingiardi, V. (2014). An empirical investigation of analytic process: contrasting a good and poor outcome case. *Psychotherapy*, 3, 270-282.
- Lingiardi, V., Gazzillo, F., & Waldron, S. (2010). An empirically supported psychoanalysis: The case of Giovanna. *Psychoanalytic Psychology*, 27, 190–218.
- Scharf, R. D., Waldron, S., Firestein, S. K., Goldberger, M. & Burton, A. (version

2010)(Unpublished). The Analytic Process Scales (APS) Coding Manual. Available from woodywald@earthlink.net.

Waldron, S., Scharf, R.D., Crouse, J., Firestein, S.K. Burton, A., & Hurst, D. (2004a).

Saying the right thing at the right time: a view through the lens of the Analytic Process Scales (APS). *Psychoanalytic Quarterly*, 73: 1079-1125.

Waldron, S., Scharf, R. D., Hurst, D., Firestein, S. K., & Burton, A. (2004b). What

happens in a psychoanalysis: A view through the lens of the Analytic Process Scales (APS). *The International Journal of Psychoanalysis*, 85, 443–466.

Waldron, S., & Helm, F. (2004). Psychodynamic features of two cognitive/behavioural and one psychodynamic treatment compared using the Analytic Process Scales. *Canadian Journal of Psychoanalysis*, 12, 346–368.

Waldron, S., Gazzillo, F., Genova, F., & Lingiardi, V. (2013). Relational and classical elements in psychoanalyses: An empirical study with case illustrations. *Psychoanalytic Psychology*, 30, 567–600.

For more information, write to Sherwood Waldron (woodywald@earthlink.net) or Francesco Gazzillo (freuwin@libero.it).

The Dynamic Interaction Scales (DIS)

The DIS (Waldron et al., 2013) are 12 5-point Likert scales (0–4) aimed at the assessment of global interactional aspects of the therapeutic process on the basis of audio-recorded and transcribed sessions. The scales were developed to address more global and relational aspects than were explored by the APS (above) because the field has shifted in a more relational direction in the years since the APS was first conceived and because an important finding using the APS had been the powerful role of therapist communication quality in contributing to short-term benefit. .

The DIS are divided into therapist scales, patient scales, and interaction scales. The therapist scales assess the degree to which the therapist is straightforward, is warmly responsive, is responsive moment-to-moment, conveys aspects of his or her subjective

experience to the patient, and is working well with, and helping the patient work with, his or her typical patterns of relating and feeling. The patient scales assess to what degree the patient flexibly shifts to and from experiencing and reflecting, shows a flexible interplay between conscious waking life and dreams, and works well with his or her typical patterns of relating and feelings. The DIS interaction scales assess the degree to which the patient experiences the therapist as empathic, to which the therapist's contribution leads to the further development of the patient's awareness, to which there is an integration of understanding of the relationship with the therapist to other relationships; and to which the engagement in the therapeutic relationship by the two parties is brought forward or experienced in an emotionally meaningful way.

The average ICC for the DIS scales assessed by three trained independent raters on 120 sessions is .68, ranging from .60 to .88 (Gazzillo et al., 2014).

Strengths. The DIS makes possible a reliable assessment of some relevant interactional and relational aspects of psychotherapy, addressing questions relevant to both the M Axis and the P Axis, such as the degree to which the patient is able to oscillate between experiencing and reflecting; and to engage himself or herself in a close working relationship that leads to greater self-understanding. The instrument has the advantage of evaluating several different aspects of the working relationship between patient and therapist, thus broadening what may be learned from the WAI (see above). Measures of these kinds of variables have been infrequent, although there is much clinical evidence to suggest their importance. The DIS is easy to apply: assessment of a 50-minute session takes about 55 minutes (listening and reading of the transcription included).

Limitations. This is a new instrument and will benefit from more extensive reliability testing and studies of convergent validity, with use by other investigators. Like other instruments based on transcripts, it is time consuming and requires training to achieve reliable ratings.

For more information, you can write to Sherwood Waldron (woodywald@earthlink.net) or Francesco Gazzillo (freuwin@libero.it).

References

- Waldron, S., Gazzillo, F., Genova, F., & Lingiardi, V. (2013). Relational and classical elements in psychoanalyses: An empirical study with case illustrations. *Psychoanalytic Psychology*, 30, 567–600.
- Gazzillo, F., Waldron, S., Genova, F., Ristucci, C., Angeloni, F., (2014). An Empirical Investigation of Analytic Process: Contrasting a Good and Poor Outcome Case. *Psychotherapy*, 51, 2, 270-282.

The Comprehensive Psychotherapy Process Scales (CPPS)

The CPPS, developed by Hilsenroth & colleagues (2005), consists of 20 items, 10 Psychodynamic-Interpersonal and 10 Cognitive Behavioral, that have been found to empirically distinguish between the two general treatment styles. Therapists or researchers rate each item on a 7-point Likert scale from 0 (*not characteristic*) to 6 (*extremely characteristic*) on the basis of transcriptions and audio or video recordings of therapy sessions.

Strengths. The CPPS is a reliable and easy-to use tool for differentiating different treatment modalities and for assessing several technical interventions that are used in psychotherapy.

Weaknesses. The CPPS does not assess the patient's contribution to the therapeutic process, and it needs an ad hoc training to be applied reliably.

References

- Hilsenroth, M. J., Blagys, M. D., Ackerman, S. J., Bonge, D. R., & Blais, M. D. (2005).

Measuring Psychodynamic-interpersonal and cognitive-behavioral techniques:
Development of A comparative psychotherapy process scale. *Psychotherapy: Theory, Research, Practice, Training*, 42, 340-356.

Table of tools described in this section

PDM AXIS	CLINICIAN REPORT TOOLS	SELF-REPORT TOOLS	PERFORMANCE BASED TOOLS
P AXIS	<p style="text-align: center;">PDC-2 PDP-2 SWAP</p> <p style="text-align: center;">KAPP (P axis functions and level of personality organization)</p> <p style="text-align: center;">OPD-2 (P axis functions and level of personality organization)</p> <p style="text-align: center;">PCRS CCRT (the conflict relationship aspect of personality patterns)</p>	<p style="text-align: center;">MMPI (personality organization and personality patterns facets)</p> <p style="text-align: center;">PAI MCMI-III (personality patterns)</p> <p style="text-align: center;">CRQ (P axis functions and level of personality organization)</p>	
M AXIS	<p style="text-align: center;">PDC-2 SWAP (PHI & RADIO)</p> <p style="text-align: center;">PFS SPC SCORS ORI AAI AAP RFS QOR PCRS CCRT DMRS DMM</p> <p style="text-align: center;">(all the instruments above may be useful for assessing several M axis capacities)</p>	<p style="text-align: center;">SIPP-118</p> <p style="text-align: center;">IIP TAS-20 (specific M axis capacities)</p> <p style="text-align: center;">CORE-OM OQ-45 SOS-10 (overall M axis level)</p>	<p style="text-align: center;">RIM</p> <p style="text-align: center;">TAT</p>

A Case Illustration: Tool Based Assessment²

Introduction to the case and referral questions

Charlotte is a 19-year-old college freshman on leave from school following a severe depressive episode in which she lapsed in basic self-care and hygiene, was unable to attend classes, and retreated to her dorm room, making minimal connection with peers. It appeared that this episode was triggered by being away from home for an extended period for the first time while facing new and unfamiliar academic and social pressures. She also was unable or unwilling to seek necessary supports on campus. At the end of the semester, when her failing and incomplete grades became apparent to her family and when the school administration alerted the family to the severity of her emotional crisis, which she had masked and minimized in weekly phone calls, she was assisted to take a medical leave of absence and return home for treatment.

Treatment involved trials of various antidepressant and anti-anxiety medications plus 1-2 times per week cognitive-behavioral therapy (CBT) aiming at behavioral activation (e.g., creating a routine of structure and activity to counter the depressive cycle created by interpersonal isolation) and to challenge maladaptive thought patterns that were leading to feelings of hopelessness and helplessness. Despite nearly six months of this outpatient plan, Charlotte's symptoms and poor functioning largely persisted. Charlotte reported that she could not tolerate different medications because of discomfiting side effects and, when side effects were absent, her conviction was that the medications were not helping. Similarly, she continually asserted therapy was "pointless . . . a waste of time" and that the therapist (a respected, well-trained senior clinician with expertise in CBT) was "useless." She frequently overslept and missed or was late to sessions (even though not scheduled early in the morning), was typically uncommunicative with her therapist, and rarely completed between-session homework assignments, even if she had reluctantly agreed to

² Contributed by Anthony Bram, PhD, ABAP. Identifying information is disguised. The author acknowledges the scoring contributions to this section by Drs. Michelle Stein (SCORS-G), Kevin Meehan (ORI), and Joseph Reynoso (ORI), as well Kiley Gottschalk and Oren Lee-Parritz.

do so. With the treatment mired in impasse, an evaluation with psychological testing was requested by both her psychiatrist and therapist, with Charlotte's parents' support, to understand more fully factors underlying and driving her symptoms, as well as what was impeding her ability to form a productive alliance with her treaters.

The older of two sisters, Charlotte grew up with her biological parents, both successful professionals with advanced degrees. There was a family history of depression, anxiety (including panic disorder), and alcoholism. Charlotte suffered mild developmental delays in fine motor skills and had unusual sensory sensitivities (tactile and auditory in particular). The delays were redressed with occupational therapy during preschool and early elementary school years. She did not have cognitive, speech, or other language delays, and there was no reported history of trauma. Temperamentally, from an early age she was viewed as shy, introverted, inhibited, anxious, and easily overstimulated. Charlotte's parents recall that, around age 3, with the birth of her sister, who had a more easy-going temperament and was better able to engage her parents and others in playful, affectionate interaction, Charlotte herself became less "cuddly" and less inclined to express and seek out affection. Though not overtly hostile and rejecting toward her baby sister, her parents recall that Charlotte was "lukewarm" to her presence in her life. In retrospect, they wondered if it was around that time that Charlotte had become more interpersonally withdrawn and emotionally shut down.

As a school-age child and adolescent Charlotte had a small group of friends who shared similar interests, but she did not seek out or make new friends readily. She excelled in and enjoyed dance and gymnastics. Charlotte was viewed as a strong but not highly motivated student until middle school, when she appeared more disorganized and had trouble completing and handing in work on time, often refusing to do so. Neuropsychological and psychoeducational testing conducted at the time revealed her to have overall Superior, albeit uneven, intellectual capacities (WISC-IV General Abilities Index at the 95th percentile, with more Average-level Working Memory and Processing Speed), with weaknesses in executive functioning and anxiety/avoidance associated with perfectionism,

which were contributing to her difficulties staying organized and completing schoolwork. Against her protests, her parents decided to place Charlotte in a small private school for the remainder of middle and high school where there was less academic and athletic pressure, the social demands more manageable, and there was better support and accommodation for her executive functioning. Her parents recall that Charlotte continues to “bristle” when the topics of her cognitive weaknesses and needs for support are explicitly broached.

Multiple efforts to help Charlotte psychotherapeutically with her anxiety did not get far. She dreaded and often refused to attend sessions, saying she hated all the “touchy-feely crap.” When her parents encouraged her to try a different therapist or modality, Charlotte had a similar response. Ultimately her parents relented and decided not to push this further. That she was completing her schoolwork, receiving solid grades, participating in dance and gymnastics, and had a couple of friends during her high school years did not create urgency regarding her need for treatment.

Recall that the present evaluation had been recommended to better understand the factors associated with her failure to adapt to the academic and social requirements of college. Treatments conceptualized and based on her seemingly accurate DSM-5 diagnoses—major depression and social anxiety—had not been engaging her, ameliorating her symptoms, or improving her functioning. Thus, for the evaluation, it was important to shift diagnostic vantage points to complement the extant DSM symptom-based perspective. Specifically, valuable alternate, complementary frameworks were offered by the combination of the PDM-2 and a treatment-centered diagnostic approach (Bram & Peebles, 2014; Peebles, 2012). Such approaches shifted the diagnostic focus toward clarifying underlying, implicit factors that create and drive manifest symptoms—that is, the role of character style and structure, ego functioning (i.e., areas of structural weakness and strength), and possible areas of intrapsychic conflict. The task was also to understand factors that affected Charlotte’s ability to form a therapeutic alliance.

Measures and methods

The evaluation comprised a combination of performance-based, self-report, and

collateral-report measures.

Performance-based measures: selected subtests of the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III), Rorschach (administered and scored with the Comprehensive System), Thematic Apperception Test (TAT; scored using the SCORS-G and Symons et al.'s [2005] Mental State Discourse Measure); Object Relations Inventory (ORI; scored on the Differentiated-Relatedness scale)

Self-report measures: Minnesota Multiphasic Personality Inventory-2 (MMPI-2), Toronto Alexithymia Scale-20 (TAS-20), Trauma History Questionnaire, Beck Depression Inventory-II (BDI-II), Depressive Experiences Questionnaire (DEQ; Blatt, D'Afflitti, & Quinlan, 1976)

Collateral-report measures: Developmental questionnaire (completed by parents), Shedler-Westen Assessment Procedure (SWAP-200; completed by referring therapist),

Diagnostic Summary based on the evaluation

Because Charlotte can present herself as a healthy, bright, athletic young woman who does not report or show great vulnerability or distress, it has likely been perplexing for her family and treaters to make sense of the severity and nature of her challenges in functioning that have necessitated her leave of absence from school. The evaluation data indicate that her emotional and behavioral difficulties are attributable to an underlying constellation of certain underdeveloped psychological capacities (i.e., structural weaknesses or deficits) and her having learned to compensate for and cope with them in ways that have become part of a maladaptive character style.

Charlotte's most significant domain of structural weakness involves *affect regulation*. This is marked by her becoming easily flooded by feelings experienced as intolerable in a way that leaves her cognitively stymied and helpless. This intensity of emotional reaction must seem confusing and alien to a young woman who is so competent, intellectually and physically or athletically. Importantly, the findings suggest that her weaknesses in affect regulation are less about her being emotionally unaware and not having the words to communicate feelings and more about her ability to sit with, tolerate, and know what to do

to manage them. She views thinking and talking about feelings not so much as a potential way to cope with and regulate them (e.g., by contextualizing them, leading to new perspectives, signaling to others how they can help) but as unnecessarily risking reevoking her noxious experience of flooding and helplessness. Despite her persistent efforts to shut out and minimize emotions, the present data show that she continues to struggle with anxiety, anger, and variable moods.

The other key area of psychological weakness identified involves underdevelopment of a *basic sense of trust in interpersonal relationships*. She has not yet internalized a stable sense that relationships and interactions with others will be the source of satisfaction and sustenance. She is leery about getting close to others, and she is vigilant about others' potentially hostile intent and has some proneness to misperceive or misinterpret in this direction. It is reasonable to speculate that she also is guarded against interpersonal closeness because of its association with sharing of emotional vulnerability which, as described above, she experiences as threatening.

Data from the evaluation indicate that she has developed entrenched, habitual ways of managing these areas of structural weakness. These entail her ongoing and automatic/reflexive efforts to (1) keep emotions out of her awareness and discourse to minimize her experience of herself as vulnerable and (2) carefully guard her personal boundaries to maintain what seems to her a safe interpersonal distance. Specifically, she has developed a *defense style* marked by constriction, minimization, rationalization ("I don't feel it's necessary" to consider and discuss emotions), externalizing (viewing distress as existing in or being caused by others), and dismissing/devaluing others viewed as potentially threatening (perhaps because they are experienced by her as holding the possibility of greater emotional involvement, positive or negative). Such a style is aimed at self-protection, preserving some sense of competence, control, and safety. It is nevertheless also maladaptive because, among other things, it limits engagement of her creativity, spontaneity, options for problem-solving, communication of needs, and satisfactions that come with deeper interpersonal/emotional engagement.

Emotional regulation

The present data indicate that Charlotte's emotional constriction and uncommunicativeness can be conceptualized as a complex combination of (1) underdevelopment/weakness in the core capacity of emotional regulation and (2) an ingrained character pattern that likely developed out of compensatory efforts at self-protection. The former is marked by Charlotte's tendency to become internally flooded by intense or complex emotions. That is to say, when feelings are stirred, she has difficulty accessing her superior cognitive abilities in the service of coping and problem-solving, is vulnerable to more confused and illogical thinking, and is more apt to misperceive situations and people (see "Reasoning and Reality Testing" section below). Anger, whether she is aware of it or not, is the source of anxiety and is particularly disruptive to such cognitive abilities. Significantly, Charlotte has little internalized sense that anxiety and states of distress and unease can be constructively resolved. This was most evident in her TAT narratives where she had difficulty bringing emotional tension states to resolution, *ending* stories instead with characters' declarations of futility, helplessness, indecision, or shutting down: "... 'why am I even trying?'" "What should I do now?" "...wondering what she should do next," and "she's going to go to bed."

Although there is some evidence that at times she is not aware of and thus might have difficulties naming the feelings impacting her, an encouraging finding is that she actually has developed a much richer emotional vocabulary than might be expected based on her history. Most notably, through her TAT responses (mostly offered spontaneously and not requiring additional prompting), she demonstrated an ability to articulate a range of pleasant emotional states (e.g., happiness, enjoyment, love, satisfaction) and also distressing ones (e.g., sadness, frustration, disappointment, fear, loneliness). This is consistent with her response to inquiry on the TAS-20 when she surprisingly endorsed the item *I am able to describe my feelings easily*. She elaborated: "I am *able to*. I just *may not want to*...in many situations... because I don't feel it's necessary." These are crucial findings that clarify that her difficulty with emotional regulation is *not* predominantly that she is oblivious about feeling states and cannot identify them (though it may be more

challenging in moments of heightened emotional intensity). It is more that, in spite of this ability, she is frightened by emotions (“affect phobia” in the terms of McCullough et al., 2003), which can acutely flood and destabilize her.

Given Charlotte’s vulnerability to the destabilizing impact of emotions, understandably she has developed strong, habitual ways of keeping feelings out of her awareness and discourse. The test data corroborate that when she *is* able to evade emotion, she actually does function better cognitively. Thus, it makes psychological sense for her to have developed this style aimed at extruding emotional awareness and expression through constriction, minimization, rationalization (“I don’t feel it’s necessary” to consider and communicate emotions), externalizing (viewing distress as existing in or being caused by others), and dismissing/devaluing those who hold that emotions are important. Striving to keep emotions out enables her to experience a greater sense of control, competence, and safe interpersonal distance. The downside, however, is that this entrenched self-protective pattern is also maladaptive, insofar as it is at the expense of her creativity, spontaneity, ability to see the big picture, communication of her needs, and depth of interpersonal engagement.

Experiences of self and others

This young woman’s Rorschach percepts of “a mask,” “armor,” “helmet,” and “shield” evoke the premium she places on self-protection and on obscuring from others her inner experience and vulnerability. Although she is certainly not uninterested in other people, Charlotte is highly cautious about emotional closeness with others. When there is a topic that engages her, she believes there to be a “point” to a specific interaction. At such times, she does not feel pressured or threatened and is able to show some reciprocal communication, perspective taking, pleasure in sharing her ideas, and sense of humor. But she is vigilant about whom she really lets in and trusts. Charlotte has a low threshold for experiencing others as intrusive, and she carefully guards her interpersonal boundaries. This vigilance is central to her difficulties making use of a therapist, as well as to her hesitation to seek out necessary support for her academic and executive functioning at

college.

Reasoning and reality testing

Testing does not reveal severe or pervasive weaknesses in Charlotte's capacities to reason logically and perceive situations accurately. Her reasoning and sense of reality are on the most solid ground when situations are more highly structured (i.e., expectations are clear and predictable, external monitoring and feedback are provided, and she is less on her own to make sense of and organize things), less emotionally demanding and intense, and have less of an interpersonal component. As suggested above (see "Emotional Regulation" section), when situations are more complex or emotionally loaded—especially if she is angry, anxious, or feeling threatened in some way—these psychological capacities can momentarily deteriorate. Specifically, under such conditions, her reasoning can become more confused and illogical, such that she is more vulnerable to reading undue hostile meaning into the situation or connecting her ideas in ways that do not make sense and lead to shaky conclusions. Under similar conditions, her perceptions and experiences of situations and people can become more distorted, often in the direction of finding threats that others in the same situation would not.

Treatment implications

The central therapeutic tasks for Charlotte are to (a) assist her to develop her capacity to tolerate and communicate her feelings and learn that there are things that she can do (cognitively, behaviorally, interpersonally) to diminish their intensity and to find solutions to handle them more successfully, (b) internalize a greater sense of trust and comfort with closeness, (c) recognize the costs of and modify her constricted, mistrustful, and avoidant defense style.

A therapeutic alliance around these aims will not be easily attained and will take time. Because of her previous difficulty accessing outpatient therapy—likely attributable to underlying mistrust and affect phobia—she will initially require a residential treatment that will mitigate against her proclivity to emotional avoidance and interpersonal withdrawal.

It is hoped that living and participating in a therapeutic milieu will offer her a level of validation, support, and feedback that would be difficult for her to access and accept from a single professional on an outpatient basis.

As regards psychotherapeutic modalities within a residential treatment program and subsequently as an outpatient, the present evaluation points to some non-mutually-exclusive options to be considered by Charlotte, her family, and treatment team. One possibility is that a therapy that weaves in biofeedback could possibly help her to link physical states of tension with particular feelings and themes and then show her concrete strategies to regulate and resolve the tension. The idea would be to make feelings more tolerable so that she can experience mastery over them that she currently lacks. Another option would be a psychotherapy group, with a stable membership and balance between a skills and process-oriented focus and run by a well-trained therapist who would be attuned to Charlotte's vulnerabilities and could play an active role in helping regulate the expectations and pace of her engagement. As is the case in the recommendation for a residential milieu, the group modality offers opportunities for peer modeling of emotional disclosure and problem-solving, as well as for feedback and confrontation about how her interpersonal style impacts others. Finally, and likely most difficult for her to accept, would be a long-term, multiple-times-per week, relationally focused psychotherapy or psychoanalysis. This too would need to be carefully paced and would necessarily move slowly for some time, especially with the therapist's recognizing and respecting her needs for constriction and guardedness. The task would be to avoid pressuring her prematurely for emotional disclosure but instead over time gradually to build trust; recognize, discuss, and affirm her strengths and interests; and model and cultivate playfulness, comfort with vulnerability, curiosity, and self-awareness. With the frequency over time, and growing safety in the relationship, there would be increasing opportunities where the here-and-now emotional interaction (positive or negative) between Charlotte and the therapist could be tolerated, reflected on, and discussed. In other words, this would be a more implicit, *in vivo* way to learn how to manage and express emotions, as well as to assist her to grapple with

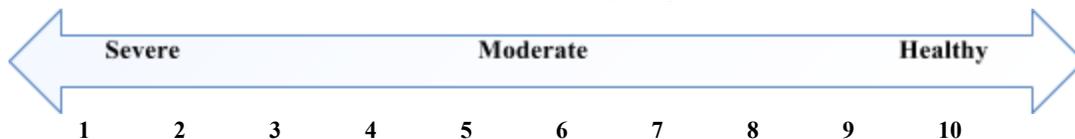
her misgivings about trust and closeness.

Evidence for case formulation from PDC and other tests

The following table, integrating scores from the PDC-2 and the other test data, illustrates how the PDC-2, applied to this case, is supplemented by several of the tools described in this section, which provide thereby a deeper and richer portrayal of her inner life, conflicts and defenses.

PDC-2 Section I: Level of Personality Organization

Consider your client's mental functions in determining the level of personality organization. Use these four mental functions to efficiently capture the level of personality organization. Rate each mental function on a scale from 1 (Severely impaired) to 10 (Healthy).



1. **Identity:** *ability to view self in complex, stable, and accurate ways* **7**

ORI, Differentiated-Relatedness scale-self description=7 on 10-pt scale: thoughts, feelings, and needs are differentiated and modulated. Increasing tolerance for and integration of disparate aspects of self

SCORS-G Identity and Self Coherence scale=4.9 on 7-pt scale (above mean for college students), identity and self-definition are not a major preoccupation

SWAP-200 –identity items not heavily weighted

2. **Object Relations:** *ability to maintain intimate, stable, and satisfying relationships* **3**

SCORS-G Capacity for Emotional Investment in Relationships=3.4 on 7-pt scale (within one standard deviation of mean)-somewhat shallow relationships

SWAP-200-elevated on Passive-Aggressive, Paranoid, and Schizoid scales, examples of heavily weighted items include “quick to assume others wish to harm or take advantage of her,” “feels misunderstood, mistreated, or victimized,” “lacks social skills”; SWAP Trait Dimensions-elevated on Hostility scale

Rorschach-lack of Cooperative movement, elevated Aggressive Content, Vulnerability to distort perceptions of people in malevolent ways

3. **Level of Defenses:** (using the guide below, select a single number) **4**

1-2: Psychotic level (delusional projection, psychotic denial, psychotic distortion)

3-5: Borderline level (splitting, projective identification, idealization/devaluation, denial, acting out)

6-8: Neurotic level (repression, reaction formation, intellectualization, displacement, undoing)

9-10: Healthy level (anticipation, self-assertion, sublimation, suppression, altruism, and humor)

SWAP-200- elevated on Paranoid scale; heavily weighted items “tends to believe her problems are caused by external factors,” “tends to become irrational when strong emotions are stirred up; may show a noticeable decline from customary level of functioning”

SWAP-200-also includes heavily weighted items “tends to express aggression in passive and indirect ways,” “has difficulty acknowledging and expressing anger,” “prefers to operate as if emotions were irrelevant or inconsequential,” “tends to be inhibited or constricted; has difficulty allowing self to acknowledge or express wishes and impulses”

ORI Differentiation-Relatedness scale Mother and Father description=5 “semi-differentiated tenuous consolidation of representations through splitting...marked oscillation between dramatically opposite qualities”

Rorschach and TAT –devaluation (e.g., “no point to this”)

4. **Reality Testing:** ability to appreciate conventional notions of what is realistic

7

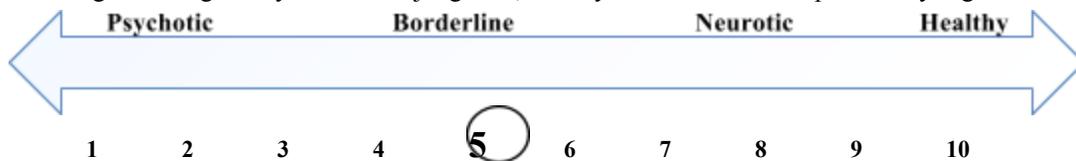
Rorschach Form *Quality* ratios are all within standard deviation of mean but evidence that reality testing can lapse under conditions of heightened affect; some vulnerability to distort perceptions of others, viewing them as more malevolent

TAT- no perceptual distortions

MMPI-2 – no report of unusual sensory or perceptual experiences

Overall Personality Organization

Considering the ratings and your clinical judgment, circle your client’s overall personality organization.



Healthy Personality- characterized by mostly 9-10 scores, life problems rarely get out of hand and enough flexibility to accommodate to challenging realities. (Use “9” for people at the high functioning neurotic level.)

Neurotic Level- characterized by mostly 6-8 scores, basically a good sense of identity, good reality testing, mostly good intimacies, fair resiliency, fair affect tolerance and regulation, rigidity and limited range of defenses and coping mechanisms, favors defenses such as repression, reaction formation, intellectualization, displacement, and undoing. (Use “6” for people who go between borderline and neurotic levels.)

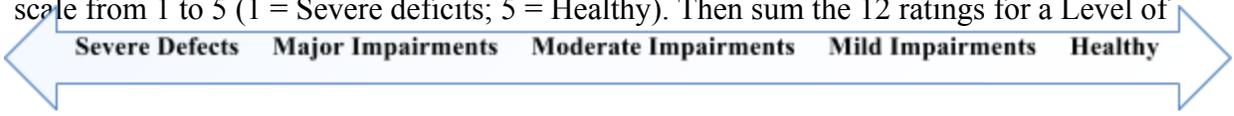
Borderline Level- characterized by mostly 3-5 scores, recurrent relational problems, difficulty with affect tolerance and regulation, poor impulse control, poor sense of identity, poor resiliency, favors defenses such as splitting, projective identification, idealization/devaluation, denial, omnipotent control, and acting out.)

Psychotic Level- characterized by mostly 1-2 scores, delusional thinking, poor reality testing and mood regulation, extreme difficulty functioning in work and relationships favors defenses such as delusional projection, psychotic denial, and psychotic distortion. (Use “3” for people who go between psychotic and borderline levels.)

(There are no sharp cutoffs between categories. Use your clinical judgment.)

Section III: Mental Functioning (M-Axis)

Rate your client's level of strength or weakness on each of the 12 mental functions below, on a scale from 1 to 5 (1 = Severe deficits; 5 = Healthy). Then sum the 12 ratings for a Level of


Severe Defects
Major Impairments
Moderate Impairments
Mild Impairments
Healthy

1
2
3
4
5

A. Cognitive and affective processes

1. Capacity for regulation, attention, and learning **3**

Previously documented weaknesses in attention and executive functioning

2. Capacity for affective range, communication, and understanding **2**

SWAP-200 heavily weighted items include "tends to be angry and hostile," "tends to avoid confiding in others for fear of betrayal," "tends to be irrational when strong emotions are stirred up," "tends to feel shamed or embarrassed," "tends to be anxious," "tends to be inhibited and constricted; has difficulty allowing self to acknowledge or express wishes and impulses"

Rorschach indicators of affective constriction

TAT difficulty resolving characters' emotional tension

3. Capacity for mentalization and reflective functioning **3**

TAT-able to identify and label mental states, Mental State Discourse measure = .35 (within one standard deviation relative to college-age norms)

Rorschach and TAT-pacing of discourse attuned to examiner

SWAP-200 heavily weighted items include "has little psychological insight into own motives, behavior, etc.; is unable to consider alternate interpretations of her experience," "tends to perceive malevolent intentions in others' words and actions," "has difficulty acknowledging or expressing anger"

B. Identity and relationships

4. Capacity for differentiation and integration (identity) **3-4**

See P-Axis above, Identity

5. Capacity for relationships and intimacy **1-2**

See P-Axis above, Object Relations

6. Self-esteem regulation and quality of internal experience **2-3**

SCORS-G Self-Esteem scale=3.9 (4="self-esteem is bland, absent, or limited")

SWAP-200- heavily weighted "tends to be self-critical; sets unrealistically high standards for self and is intolerant of own human defects"

Rorschach-no structural indicators of grandiosity

Symptom/Concern: *Social Anxiety* Level: **2-3**

Section V: Cultural, Contextual and Other Relevant Considerations

History of weaknesses in executive function ; High-achieving family members; competitiveness with younger sibling

References

- Blatt, S. J., D’Afflitti, J.,P. & Quinlan, D. M. (1976). Experiences of depression in normal young adults. *Journal of Personality Assessment*, 85, 383-389.
- Bram, A.D., & Peebles, M.J. (2014). *Psychological testing that matters: Creating a road map for effective treatment*. Washington, DC: American Psychological Association.
- McCullough, L., Kuhn, N., Andrews, S., Kaplan, A., Wolf, J., & Hurley, C.L. (2003). *Treating affect phobia: A manual of short-term dynamic psychotherapy*. New York, NY: Guilford.
- Peebles, M.J. (2012). *Beginnings: The art and science of planning psychotherapy* (2nd ed.). New York, NY: Routledge.
- Symons, D., Peterson, C., Slaughter, V., Roche, J., & Doyle, E. (2005). Theory of mind and mental state discourse during book reading and story-telling tasks. *British Journal of Developmental Psychology*, 23, 81-102.