MANUAL FOR THE

STRUCTURED INTERVIEW OF PERSONALITY ORGANIZATION-REVISED

(STIPO-R)

By

John F. Clarkin, Eve Caligor, Barry Stern, & Otto F. Kernberg

July, 2019

Revised July 2021

This manual for the Structured Interview of Personality Organization-Revised (STIPO-R) is composed of two sections. Section 1 provides a history, background and rationale, and review of psychometric properties and utilization of the instrument. Section 2 provides an interview guide to assist in the actual administration of the instrument.

SECTION 1: INTRODUCTION AND REVIEW OF STIPO-R

History and Background of the STIPO and STIPO-R

The STIPO (Clarkin, Caligor, Stern, & Kernberg, 2004), and its revision, the STIPO-R (Clarkin, Caligor, Stern, & Kernberg, 2016) are semi-structured interviews constructed to evaluate the structural domains of personality functioning that are central to understanding the individual from an object relations model of personality and personality pathology (Kernberg, 1984; Kernberg & Caligor, 2005). The STIPO and the STIPO-R provide the clinician and researcher with dimensional scores on key domains of personality functioning. The severity of dysfunction in each domain can be used by the clinician for treatment planning, and by the researcher for selection of subjects and measurement of change in relation to treatment interventions.

Object Relations Orientation to Personality Pathology

Kernberg and colleagues at the Personality Disorders Institute have articulated a model of personality pathology based in contemporary object relations theory (Kernberg & Caligor, 2005; Caligor & Clarkin, 2010; Caligor, Kernberg, Clarkin, & Yeomans, 2018). This approach combines a dimensional view of severity of personality pathology with a categorical or prototypic classification based on descriptive phenomenology consistent with many of the personality syndromes of DSM-5 (APA, 2013). Thus, the STIPO and STIPO-R provide both severity scores on domains of functioning, and profiles of scores in the domains indicating closeness/distance to prototypic descriptions of neurotic, high level borderline, middle level borderline, and low-level borderline personality organization. Level of personality organization has important prognostic implications and can be used to guide differential psychotherapeutic treatment planning (Caligor, Kernberg, Clarkin, & Yeomans, 2018).

Core concept of identity. Kernberg's object relations model of personality pathology is organized around the core concept of "identity." The universe of personality disorders is divided into those characterized by consolidated identity and those characterized by pathology of identity formation (sometimes referred to as the syndrome of "identity diffusion").

The less severe (neurotic) level of personality organization (NPO), along with the normal personality, is characterized by a consolidated identity associated with an experience of self and of others that is stable, well differentiated, complex, realistic and coherent. The neurotic level of personality organization is distinguished from the normal personality on the basis of rigidity of personality functioning. Whereas the individual with normal personality organization is able to flexibly and adaptively manage external stressors and internal conflicts, the individual with neurotic personality organization tends to rely on rigid and to some degree maladaptive responses, reflecting the impact of repression-based defenses on psychological functioning. As in the normal personality, individuals organized at a neurotic level have the capacity for full, deep and mutual relationships, though individuals in the NPO spectrum may have difficulty combining intimate relations with sexuality. Moral functioning is consistent and fully internalized in the neurotic personality, but may be excessively rigid, leading to a propensity to excessive self-criticism.

Identity diffusion is a major characteristic of the borderline level of personality organization (BPO). Poorly consolidated identity is associated with an experience of self and others that is unstable, superficial, poorly differentiated, polarized ("black and white"), distorted and discontinuous. Splitting-based defenses (e.g., splitting, idealization/devaluation, projective identification, denial) are responsible for maintaining a fragmented and poorly integrated experience of self and others that color the subjectivity of the individual with poorly integrated identity. In contrast, consolidated identity in the neurotic personality disorders is associated with the predominance of repression-based and mature defensive operations. Individuals organized at a borderline level of personality organization are distinguished from those with atypical psychotic disorders by virtue of having intact reality testing. However, clinically significant pathology of identity formation is associated with deficits in social reality testing, the ability to accurately infer the motivations and internal states of others and to accurately read social cues. These deficits are associated with some impairment in accurate perception of others in individuals organized at a borderline level of personality organization. In contrast, social reality

testing is highly developed in individuals organized at a neurotic level, as well as in the normal personality.

The borderline level of personality organization, characterized by identity pathology, the predominance of splitting-based defenses and deficits in social reality testing, covers a relatively broad spectrum of personality pathology. At the higher end of the BPO spectrum, patients have some capacity for dependent, albeit troubled, relationships, generally have relatively intact or only minor pathology of moral functioning and are not overtly aggressive in most settings. In contrast, individuals at the lower end of the BPO spectrum have severe pathology of object relations, clinically significant deficits in moral functioning, and are overly aggressive, while those in the middle BPO spectrum have moderate pathology of object relations, variable moral functioning and demonstrate pathology of aggression less severe than is characteristic of the low BPO spectrum. Whereas individuals in the high BPO group have a relatively favorable prognosis in structured psychodynamic treatments, and those in the middle BPO group have a fair prognosis, those in the low BPO group are far more challenging to treat and have a more guarded prognosis, even in exploratory psychodynamic treatments with established parameters designed to ensure that the patient's difficulties are expressed in the treatment situation and to limit self-destructive and treatment-interfering behaviors outside the consulting room.

Determination of level of personality organization is essential to guiding differential treatment planning. Psychodynamic intervention with high level personality functioning (neurotic organization) is constructed differently (Caligor, Kernberg, & Clarkin, 2007; Caligor, Kernberg, Clarkin, & Yeomans, 2018) than intervention with patients at borderline levels of organization (Yeomans, Clarkin, & Kernberg, 2015) (see Table 1). Individuals organized at a neurotic level of personality organization have a very favorable prognosis and can benefit from relatively unstructured psychodynamic treatments. These patients typically do not have difficulty establishing and maintaining a therapeutic alliance, and transference distortions tend to be slowly developing, consistent, and subtle. In contrast, individuals organized at a borderline level, particularly those in the low borderline spectrum, require a highly structured treatment setting as described above. These individuals have great difficulty establishing and maintaining a therapeutic alliance; transference distortions develop rapidly, and are highly affectively charged and extreme, often leading to disruption of the treatment.

Table 1. Treatment Differences Related to Level of Personality Organization

NEUROTIC PERSONALITY ORGANIZATION	BORDERLINE PERSONALITY ORGANIZATION
Use of treatment frame	Treatment frame includes a carefully articulated treatment contract
Therapist operates from a stance of therapeutic neutrality	Therapist deviations from therapeutic neutrality are used in certain crises
Therapeutic techniques of clarification, confrontation, interpretation	More extensive use of clarification and then confrontation to set the stage for interpretation
Focus on present, related to past	Focus on the present

The origins of the STIPO: "The Structural Interview"

As part of articulating an object relations approach to personality pathology, Kernberg (1984) described the structural interview, a clinical interview designed to evaluate not only the patients' symptoms and areas of difficulty, but also the level of personality organization. At that time, Kernberg conceived of the structural interview in the context of existing psychodynamic interviews. A number of analytic authors had constructed modified psychiatric interviews that concentrated on the patient-therapist interaction as a major source of information (Whitehorn, 1944; Powdermaker, 1948; Fromm-Reichmann, 1950; Sullivan, 1954). Deutsch (1949) advocated interviewing that would reveal the unconscious connections between current difficulties and the patient's past. MacKinnon and Michels (1971; MacKinnon, Michels, & Buckley, 2006) described an evaluation that uses the patient-therapist interaction to reveal character patterns useful for diagnosis. Kernberg's structural interview was an organized extension of these procedures. The interview focuses on the patient's conflicts thereby creating tension such that the patient's predominant defensive and structural organization of mental functioning emerges, and the structural diagnosis of personality organization can be made.

The sequence of the structural interview proceeds through three phases. The initial phase invites the patient to discuss major difficulties, symptoms, and reasons for seeking treatment. The middle phase focuses on potential pathological personality traits, and difficulties in interpersonal relations and perceived interpersonal needs. In the termination phase, the interviewer provides an opportunity for the patient to ask questions, and for the interviewer to evaluate the patient's motivation for continuation of the diagnostic process and treatment.

The yield of the structural interview is an assessment of both symptoms and the level of personality organization, characterized by levels of organization from identity consolidation with difficulties in object relations, to high level borderline personality organization with identity diffusion, to low level personality organization with identity diffusion combined with aggression, severe pathology of object relations, and deficits in moral functioning. The yield of the structural interview depends upon the clinical acumen and skill of the interviewer. The interviewer must make sophisticated decisions about which areas of the patients' functioning to evaluate in detail. The detailed examination of the patient's relations with others provides the interviewer with an opportunity to observe the patient's functioning in a tense situation. There is no scoring system, and the interviewer must make subjective judgments about the patient's degree of personality pathology and level of personality organization. With its dependence on interviewer skill, flexibility in interview questions, and absence of an objective scoring system, it is difficult to ascertain reliability among different interviewers in terms of focus and diagnostic conclusions. These shortcomings of a sophisticated clinical interview led to the construction of the STIPO.

Need for a Semi-Structured Interview

The generation of the semi-structured interview (STIPO and STIPO-R) provides standard questions, follow-up probes, and scoring guidelines to ensure reliability in the assessment[1]. What the STIPO loses in the subtle interview maneuvers of an experienced clinician, the STIPO gains in psychometric properties. With its structured questions, and equally structured probes following vague or imprecise patient answers, and a structured scoring system, the STIPO lends itself to investigation of its reliable administration and scoring. The standardization of procedure and scoring in the STIPO-R enhances its usefulness in the teaching of personality assessment, and it provides a vocabulary that clinicians can use to clearly communicate complicated clinical constructs to each other and to those not involved in object relations theory.

Domains of the STIPO and STIPO-R

A key question in the generation of any assessment instrument concerns the selection of a limited number of domains of functioning that are crucial to the adjustment of the individual (see

[1] We wish to acknowledge the important contributions by multiple colleagues to the development of the STIPO and STIPO-R. Armand Loranger, the author of the IPDE, was a consultant who helped guide the structure of the STIPO. Mark Lenzenweger has provided valuable scoring and design advice. Susanne Hörz utilized the STIPO in her dissertation and stimulated the profile analysis of the STIPO. She and her colleague Stephan Doering have advanced the German version of the instrument. Emanuele Preti was instrumental in the transition of STIPO to STIPO-R, and has initiated important research with the instrument in an Italian version.

Clarkin, 2013). Object relations theory guides the assessor to two central phenomena in understanding the individual patient: the structure and the organization of the personality. The focus of our approach is on structural, functional domains rather than solely on the assessment of difficulties and symptoms. Object relations theory guides the selection of the domains to be assessed in the STIPO. As generated by clinical experience and psychoanalytic theory, the basic structures of personality are identified as: 1) identity, i.e., an integrated concept of the self and an integrated concept of significant others, 2) a capacity for a broad spectrum of affect dispositions that are complex and well-modulated, 3) an integrated and mature system of internalized values, and finally, and 4) an appropriate management of sexual, dependent, and aggressive motivations which are experienced subjectively as needs, impulses, wishes, and fears (Kernberg & Caligor, 2005). Following from this theoretical position, the STIPO was focused on six domains: Identity, Defenses, Quality of Object Relations, Coping, Aggression and Moral Values. With the accumulated experience with the longer STIPO, we have incorporated five key domains in the STIPO-R: Identity, Defenses, Quality of Object Relations, Aggression, and Moral Values (see Table 2).

The Identity domain is measured by questions concerning the individual's capacity to invest and be involved in studies and/or work and professional life, and recreation. The individual's sense or representation of self and of others is examined. The domain of Quality of Object Relations involves the assessment of the individual's interpersonal relations, intimate relations and sexuality, and the internal or mental model of relationships. The domain of Defenses provides an assessment of both more advanced and mature defenses, and more primitive defenses such as splitting. The domain of Aggression focuses on both aggression toward the self and aggression toward others. Finally, the domain of Moral Values or moral functioning is an examination of the individual's capacity for guilt and adherence to common norms of interpersonal behavior.

STIPO Compared to Similar Instruments

Possibly the nearest clinical interview and scoring system to the STIPO is the Clinical Diagnostic Interview (CDI; Westen & Muderrisoglu, 2003) that focuses on reasons for treatment, symptoms, and interpersonal interaction patterns. It is a systematic diagnostic interview that can be administered in two and one-half hours. The interview yields the clinical information necessary to utilize the Shedler-Westen Assessment Procedure (SWAP-200; Shedler & Westen,

2007) reliably. The SWAP-200 is an assessment instrument that consists of 200 statements that may describe a patient very well, somewhat, or not at all. The statements reflect content capturing personality traits in non-clinical populations, and interpersonal pathology consistent with personality disorder (coping, defense, and affect-regulatory mechanisms) as well as symptoms such as anxiety and depression. Utilizing the information from the CDI, the clinician describes the patient with the 200 SWAP items based on a Q-sort method which requires the clinician to distribute the 200 items into a fixed distribution, i.e., a set number that are least and most descriptive of the individual (Shedler, 2015). The SWAP distribution provides the clinician with dimensional scores for each of the personality disorders described in DSM. In addition, a narrative case description is generated that can be used for case conceptualization and treatment planning.

The Operationalized Psychodynamic diagnosis (OPD-2; OPD Task Force 2008), devised by a group of psychoanalytic clinicians in Germany, Austria, and Switzerland is an instrument consisting of four psychodynamic axes as well as the ICD-10 as a fifth axis: 1) Experience of Illness and Prerequisites for Treatment, 2) Interpersonal Relations, 3) Conflicts, 4) Psychic Structure and, 5) Psychic and Psychosomatic Disorders (ICD-10 diagnoses). The axis that most closely relates to the STIPO is the fourth axis, which comprises dimensions of self and other representation, attachment, affect differentiation or impulse regulation. OPD-2 was developed to assess all levels of personality pathology, whereas the STIPO focuses specifically on the nuances and levels of personality organization. As hypothesized, the STIPO level of personality organization was significantly related with the OPD axis 4 total score ($r=.68$; $p<.001$) (Doering, Burgmer, Heuft, et al., 2013).

Transition from STIPO to STIPO-R

The STIPO-R is a revision of the original STIPO, undertaken to both shorten the longer STIPO to enhance its research and clinical usage, and to modify items that had less than desirable psychometric properties. In addition, our clinical experience motivated us to amplify the items in the original STIPO concerning narcissistic pathology into a full Narcissism scale.

Scope of the STIPO-R

Content

The STIPO-R contains 55 items covering five domains of functioning: 1) Identity, 2) Object Relations, 3) Defenses, 4) Aggression, and 5) Moral Values. Three of the domains have

ratings on important subdomains (see Table 2). From items embedded in the other domains, the STIPO-R also has scoring for a Narcissism dimension.

Table 2. STIPO-R Domains and Subdomains

Domain	**Subdomain**
Identity 15 items	Capacity to invest in work/studies and recreation Sense of self Sense of others
Object Relations 15 items	Interpersonal relations Intimate relationships and sexuality Internal working model of relationships
Defenses 10 items	Lower-level, primitive defenses Higher-level defenses
Aggression 9 items	Self-directed aggression Other-directed aggression
Moral Values 6 items	Experience of guilt; moral and immoral behavior

Format

The format of the STIPO-R is carefully modeled on the International Personality Disorder Examination (IPDE; Loranger, 1999) constructed by our Cornell colleague, Dr. Armand Loranger. Dr. Loranger served as a consultant to the construction of the STIPO. STIPO-R utilizes standard questions, and additional probes that can be used when the answers are not clear or detailed enough to rate.

Scorning System

The standardized format and scoring system allow the interviewer to rate the subject's responses (0, 1 or 2) at the individual item level as the interview proceeds. As with the IPDE, the interviewer is encouraged to use not only information from the subject but also any additional information from ancillary sources (e.g., family members, former therapists) that may be available given the constraints of the interviewing situation, to arrive at the most accurate item ratings. Once the interview is completed, the scores at the individual item level are summed within each domain to give a total domain score. (An alternative method is to compute a mean rating from the 0-1-2 item scores across each domain.) This dimensional rating provides an indication of the total

pathology in each domain. In order to directly compare the scores among domains, the total dimensional rating can be transformed into a percentage score.

In addition to the dimensional sum scores for each domain, the interviewer is invited to make an overall clinical rating (ranging from a 1-5 score) for each domain. This overall domain rating allows the interviewer to use clinical judgment and impression about the subject that may deviate somewhat from the item total dimensional rating for each domain.

The two rating systems complement each other. The item-based rating system adheres closely to the individual item responses, whereas the 5-point rating system allows the interviewer to utilize his clinical impression, allotting greater or lesser weight to items in the scale or subscale based on his clinical impression of pervasiveness or severity, and/or adjusting the rating based on factors (non-verbal, interpersonal) that he feels are clinical significant and relevant to the domain being assessed. For both rating procedures we have found satisfactory inter-rater reliability (Stern et al., 2010; Hörz et al., 2009). The scores in table 3 demonstrate the use of these clinically oriented ratings.

Table 3 (updated 2021): STIPO Dimensional Scores and Level of Personality Organization

STIPO Dimensional Ratings	Normal PO	Neurotic-level PO 1	Neurotic-level PO 2	High-level BPO	Middle-level BPO	Low-level BPO
Identity	1	1	2	3	4	4 or 5
					Range 3-5	
QOR	1	2	3	3	4	5
		Range 1-2	*Range 2-3*	*Range 2-4*	*Range 3-5*	*Range 4-5*
Defense	1	2	2	3	4	4 or 5
		Range 2-3			*Range 4-5*	
Aggression	1	2	2	3	4	4 or 5
		Range 1-3	*Range 1-3*	*Range 2-3*	*Range 3-4*	
Moral Values	1	2	3	3	3	4 or 5
		Range 1-2	*Range 2-3*	*Range 2-3*	*Range 3-4*	

Using either the dimensional summary score for each domain or the clinical 5-point ratings, the interviewer can construct a profile of personality organization of the subject, based on the five domains of interest. Patients can be classified as falling into normal, neurotic, or borderline range of organization. Based on the STIPO-R dimensional ratings, this categorization can be made, differentiating normal, neurotic, and borderline personality organization, which is differentiated into three levels according to severity: high, mid, and low BPO (Hörz, Stern, Caligor, et al, 2009). Subjects falling into normal and neurotic group have consolidated identity; show no use of

primitive defenses or disturbance in reality testing. Patients falling into Neurotic group have some degree of superficiality in sense of self and/or others and might show some use of primitive defenses. Patients located at borderline level of personality organization range from high to mid to low, with an increase in identity diffusion, the use of primitive defense mechanisms, overt manifestations of aggression, disturbance of object relations increase, and diminished use of internal standards of morality. The following prototypical profiles aid the assignment of levels of personality organization based on the dimensional scores.

NORMAL

Identity:	1	Consolidated
Object Relations:	1	Stable, complex, and enduring; able to integrate tender and erotic feelings
Defenses:	1	Healthy defenses predominate; No evidence of primitive defenses
Aggression:	1	No primitive aggression; control and modulation of aggression
Moral values:	1	Consistent and flexible; no antisocial behavior

NEUROTIC 1

Identity:	1	Consolidated
Object relations:	2	Stable, complex, and enduring; difficulty integrating tender and erotic feelings
Defenses:	1	No evidence of primitive defenses, Repression-based and mature defenses predominate; some rigidity.
Aggression:	2	No primitive aggression; some evidence faulty modulation of aggression (e.g., minor self-neglect or occasional verbal outbursts)
Moral values:	2	Overly harsh and/or inflexible but fully organized and internalized; no antisocial behavior

NEUROTIC 2

Identity:	2	Consolidated; somewhat superficial sense of self and/or others
Object relations:	2, 3	Somewhat superficial, but enduring; some limitation in capacity for empathy; difficulty integrating tender and erotic feelings

Defenses:	2	Endorsement of primitive defenses is rare; Repression-based defenses predominate; some rigidity.
Aggression:	2	No primitive aggression; evidence of faulty modulation of aggression (e.g., minor self-destructive behaviors or controlling interpersonal style)
Moral values:	2, 3	Organized and internalized but variable; self-critical attitudes and demanding standards may co-exist with disavowal of exploitative or minor self-destructive behaviors

BORDERLINE 1

Identity:	3	Identity diffusion, mild
Object relations:	3	Split and/or superficial but with some degree of stability and integration, especially in non-conflictual domains
Defenses:	3	Combined use of splitting-based and repression-based defenses, significant rigidity and compromised adaptation.
Aggression:	3	Primitive aggression; aggressive behaviors largely self-directed
Moral values:	2, 3	Variable

BORDERLINE 2

Identity:	4	Identity diffusion, moderate
Object Relations:	4	Superficial and based on need-fulfillment; empathy impaired; little Ability to sustain interest over time; widely split and unstable
Defenses:	4	Predominance of primitive defenses with significant impairment; severe rigidity and grossly maladaptive defensive strategies
Aggression:	3, 4	Primitive aggression; aggressive behaviors directed against others +/- against self
Moral values:	2, 3,	Variable, generally poorly integrated and poorly internalized

BORDERLINE 3

Identity:	5	Identity diffusion, severe
Object Relations:	5	Based entirely on need-fulfillment, no empathy and no capacity to sustain interest in others
Defenses:	5	Constant use of primitive defenses; extreme rigidity and failure of adaptation
Aggression:	5	Primitive aggression with dangerous, aggressive behaviors directed towards self and/or others
Moral values:	5	No organized moral value system; antisocial behavior

Appropriate Subjects

All patients applying for treatment can be assessed with the STIPO-R, which provides an overall picture of the level of personality organization that influences any treatment, including those focused almost entirely on symptom constellations such as anxiety and depression. However, the STIPO-R is most relevant in clinical situations in which the patient is suspected of having personality pathology that will influence symptom treatment, or those whose treatment will focus primarily on personality disorder of various levels of severity.

Examiner Qualifications and Training

Prior training in psychodynamic concepts central to the instrument, and clinical experience with patients demonstrating various levels of severity of personality pathology are prerequisites for STIPO-R interviewers. The interviewer must be trained to use the probes to obtain ratable material from the patient. Training to reliability of scoring involves viewing of videotaped STIPO-R interviews, and accomplishment of ratings in agreement with standards.

Reliability and Validity of the STIPO

The data presented here are related to the STIPO. The psychometric properties of the STIPO-R are currently under evaluation. Preliminary unpublished data show acceptable reliability and good convergence with external measures of personality functioning.

Reliability

English, German, and Italian versions of the STIPO have been developed concurrently and have demonstrated good inter-rater reliability. Intraclass correlation coefficients (ICC) ranged from .84-.97 in the English version (Stern et al, 2010), from .89-1.0 in the German version (Doering et al, 2013), and from .82-.97 in the Italian version (Preti et al, 2012).

Validity

The STIPO domains show internal consistency across studies. Cronbach's alpha for STIPO domains of Identity (.86) and Primitive Defenses (.85) were high, whereas the shorter Reality Testing domain (.69) was on the boarder of acceptability (Stern, Caligor, Clarkin, et al, 2010). In a study using the German language version of the STIPO (Doering et al, 2013), Cronbach's alpha ranged from .93 for Identity to .69 for Reality Testing, with .97 for the total score.

STIPO domains of Identity and Primitive Defenses were closely related to personality disorder symptom counts as assessed by the Schedule of Nonadaptive and Adaptive Personality (SNAP; Clark, 1993), to measures of aggression, and to levels of positive and negative affect. In another

study (Doering, Burgmer, Heuft, et al, 2013), significant correlations were found between the STIPO Primitive Defenses and the primitive defenses scale of the self-report Borderline Personality Inventory (Leichsenring, 1997).

Preti et al, (2012) found associations between the STIPO identity scale with measures of stability of self-image and the capacity of pursuing goals. The STIPO Defenses domain was associated with an external measure of primitive defenses, and another measure of lack of self-control and emotional instability (SIPP-118; Verheul et al, 2008). All of the STIPO domains discriminated between clinical and nonclinical subjects.

The STIPO demonstrates good construct validity in reference to DSM personality diagnoses. Patients with DSM personality disorder were found to be on a lower level of personality organization in all domains compared to patients without personality disorder (Baumer, 2010; Doering, et al, 2013). In a study of patients with chronic pain, there was a significant correlation between personality organization on the STIPO and the number of SCID-II diagnoses (Fischer-Kern, et al, 2011). Likewise, a very close but not complete association was found between STIPO structural diagnoses and DSM personality pathology in a sample of patients with opiate addiction (Rentrop Zilker, Lederle, Birkhofer, & Hörz, 2014). There is a significant association between STIPO structural characteristics and DSM diagnoses, but STIPO domains were able to identify treatment dropout among dual-diagnosis patients more effectively than personality disorder diagnoses (Preti, Rottoli, Dainese et al, 2015).

Clinical application of the STIPO: Measuring severity of personality pathology

The STIPO can be used as clinical tool to assess levels of severity of personality pathology across normal, neurotic, and high- and low-level borderline personality organization. In a study using the English version of the STIPO, based on the domain ratings of the STIPO, a prototypical profile of BPO was developed and tested in its ability to discriminate between BPO and non- BPO (Hörz, 2007). The presence of severe identity diffusion, use of primitive defenses as well as disturbed object relations, along with overall maintained reality testing differentiated between patients located at low BPO and non-BPO. Individuals with ratings that were close to a prototypical profile of BPO, consisting of ratings of 3 or higher in the domains “Sense of Self” and “Sense of others”, 4 or higher in “Object Relations” and “Primitive Defenses”, showed more pathology in variables closely associated with borderline pathology, for example negative affect and aggression. Similarly, an inverse relation between the profiles of individuals with BPO-

prototypical ratings and variables of positive affect was found, e.g. serenity. In addition, evidence of poorly integrated aggression and the deterioration of moral values were helpful in differentiating between higher level and lower level BPO (Stern et al., 2010).

In a treatment study examining 104 patients with Borderline Personality Disorder (BPD), the STIPO was employed and compared to results from the SCID-I and SCID-II as well as indicators of clinical severity of the disorder (suicide attempts, self-injurious behavior, service utilization) (Doering et al, 2010). Specific patterns were found, demonstrating the ability of the STIPO to assess levels of severity. The patient group with one or more comorbid DSM-personality disorders showed more pathology in the STIPO domains and overall level of personality organization than the patient group with the sole diagnosis of BPD (e.g. Identity: $M = 3.88$ *vs.* $M = 3.59$, $t = -2.13$, $p < .04$). Similar results were found for individuals with at least one suicide attempt versus no suicide attempts, and also for patients with a history of emergency room visit versus those without emergency room visits. Moreover, correlational analyses showed that several indices of personality pathology, for example the number of BPD-criteria, was meaningfully associated with more pathology in the STIPO domains of Identity, Primitive Defenses, Coping, Aggression and with the overall level of personality organization ($r = .29$, $p < .01$). In sum, these results demonstrate the clinical usefulness of the STIPO in that patients with clinically more severe disorder revealed a more impaired level of personality organization (Hörz et al., 2017).

Clinical application of the STIPO: Using the STIPO as a measure of change

The usefulness of the STIPO as a measure to assess changes in personality organization was examined in an RCT comparing the efficacy of Transference-Focused Psychotherapy (TFP) to treatment by experienced community psychotherapists in a sample of 104 BPD patients (Doering et al, 2010). The time frame in the STIPO usually refers to the prior five years. However, in order to assess changes within one year of treatment the investigators chose the last month as the time frame for the second STIPO interview in this study. Using this measure, significant changes after one year of psychotherapy were found at the level of personality organization. In this analysis, the overall level of personality served as the outcome variable, using the STIPO levels of personality organization on a 6-point categorical scale, ranging from normal (1) to Borderline 3 (6). In both treatment groups, the mean for the level of personality organization pathology decreased after one year of therapy. This was the case both for patients in TFP (pre: M=5.00, SD=0.56; post: M=4.46, SD=0.67; d=1.0, p<.001) and for patients in the

community psychotherapist group (pre: *M*=4.77, *SD*=0.58; post: *M*=4.62, *SD*=0.53; *d*=0.3, *p*=.004), with a significant superiority for the TFP group *(F*=12.136; *df*=1, 101; *p*=.001) (Doering et al., 2010). A more detailed analysis of changes in the individual STIPO domains is currently ongoing.

Use of the STIPO for Treatment Planning and Change

The diagnosis of personality disorders by categories or types without taking into consideration the dimension of severity of dysfunction represents a serious lack in DSM-5, diminishing the utility of DSM-based diagnoses for treatment planning. One unfortunate result of this deficiency in DSM diagnosis is that existing psychotherapy treatment trials do not take into account the severity of the personality dysfunction in data analysis. The ability to use the five STIPO-R domains of functioning to match prototypic models of neurotic personality functioning, and various levels of borderline personality organization and functioning, i.e., to speak to the issue of severity within diagnostic categories, will, we hope, represent an improvement over the DSM system and provide a tool for studying the impact of clinical severity as it relates to diagnosis, and treatment process and outcome.

DSM 5, PDM2, and the STIPO-R

The DSM-5 description of personality disorder is based on lists of symptoms, traits and problematic behaviors. This list adheres closely to reportable and observable behaviors with the intent of ensuring reliability of assessment. This symptom-oriented description/assessment of personality disorders is not guided by a theory of personality or an articulated theory of the personality disorders.

In contrast, the STIPO and STIPO-R are theory driven in their conceptualization and dimensional profiles. The advantage of a theory driven assessment is that the theory provides a guide for efficient use of assessment time. A theory guided assessment also ensures that in the limited time, one assesses essential areas of personality and personality disorder functioning. For example, current theories of personality indicate that the major areas to consider are cognitive-affective units, behavior, and the person's unique pattern of relating to and seeking out certain environments. A theory guided assessment of essential areas of personality functioning can subsequently and logically lead to focused interventions on the areas of dysfunction.

The yield or product from the STIPO-R can be compared to that provided by semi-structured interviews of personality pathology such as the SCID II. The yield or product from the

SCID II is a diagnosis of one or more of the ten personality disorders as described by DSM. There is little theoretical basis behind the personality disorders in DSM-5, and the categories as described do not hold up to empirical investigation. In contrast, the yield of a STIPO-R interview is dimensional ratings of domains of personality functioning. Scores on these domains provide a profile of the patients' functioning that range from areas of adequate to inadequate functioning. The resulting profile can be used to assist the interviewer to assess the closeness of the patient to prototypic descriptions of patients at a neurotic, high- or low-level borderline organization (Hörz, 2007). This approach to personality assessment is consistent with object relations theory and is also consistent with the direction that the DSM-5 is taking with the Section III approach to the dimensional assessment and diagnosis of personality pathology. In fact, the STIPO-R is a reliable tool for obtaining patient information that can be used to make the level of personality functioning ratings in DSM-5, Section III (Preti, Di Pierro, Costantini, et al, 2018).

The Psychodynamic Diagnostic Manual-2 (Lingiardi & McWilliams, 2017) is an effort to bring diagnosis and related treatment planning closer to a theoretically coherent view of personality functioning/dysfunctioning and related symptom disorders. The object relations orientation to personality functioning is explicitly referenced in this system, and the STIPO is noted as a key instrument related to the clinical assessment of patients.

Translations of the STIPO and STIPO-R

We have encouraged colleagues at other sites to translate the STIPO into their local languages. There are established versions of the STIPO in English (Stern et al, 2010), German (Doering et al, 2013), and Italian (Preti, Prunas, Sarno, & De Panfilis, 2012). Researchers are working on versions of the STIPO-R in Poland, China, Turkey, Hungary, Czec Rebublic, Russia, Argentina, and Brazil.

Limitations of the STIPO-R

Like all interviews, the STIPO-R is limited by the honesty and ability of the subject to provide detailed and accurate information. However, unlike self-report questionnaires, the interview format provides an opportunity for the interviewer to probe and obtain further amplification from the subject and from significant others such as family and former therapists.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. (5th ed.). Washington, DC: Author.

Caligor, E., & Clarkin, J.F. (2010). An object relations model of personality and personality pathology. In: J.F. Clarkin, P. Fonagy, & G. Gabbard (Eds.), *Psychodynamic psychotherapy for personality disorders: A clinical handbook* (3-36). Washington, DC: American Psychiatric Publishing.

Caligor, E., Kernberg, O.F., Clarkin, J.F., & Yeomans, F.E. (2018). *Psychodynamic therapy for personality pathology: Treating self and interpersonal functioning.* Washington, DC: American Psychiatric Association Publishing.

Clark, L.A. (1993). *Manual for the Schedule for Non-adaptive and Adaptive Personality (SNAP).* Minneapolis, MN: University of Minnesota Press.

Clarkin J.F., Caligor E., Stern B., & Kernberg O.F. (2004). *The Structured Interview for Personality Organization (STIPO).* Unpublished manuscript, Department of Psychiatry, Weill Cornell Medical College, NY, New York.

Clarkin, J.F. (2013). The search for critical dimensions of personality pathology to inform diagnostic assessment and treatment planning: A commentary on Hopwood et al. *Journal of Personality Disorders, 27(3),* 303-310.

Clarkin J.F., Caligor E., Stern B., & Kernberg O.F. (2016). *The Structured Interview for Personality Organization-Revised (STIPO-R).* Unpublished manuscript, Department of Psychiatry, Weill Cornell Medical College, NY, New York. Download from www.borderlinedisorders.com.

Deutsch, F. (1944). *Applied psychoanalysis*. New York, NY: Grune & Stratton.

Di Pierro, R., Preti, E., Vurro, N., & Madeddu, F. (2014). Dimensions of personality structure among patients with substance use disorders and co-occurring personality disorders: A comparison with psychiatric outpatients and healthy controls. *Comprehensive Psychiatry, 55(6),* 1398-1404.

Doering, S., Hörz, S., Rentrop, M., Fischer-Kern, M., Schuster, P., Benecke, C., Buchheim, A., Martius, P., & Buchheim, P. (2010). Transference-focused psychotherapy versus treatment by community psychotherapists for borderline personality disorder: Randomized controlled trial. *British Journal of Psychiatry*, *196(5),* 389-395.

Doering, S., Burgmer, M., Heuft, G., Menke, D., Baumer, B., Lubking, M., Feldmann, M., Hörz, S., & Schneider, G. (2013). Reliability and validity of the German version of the Structured Interview of Personality Organization (STIPO). *BMC Psychiatry, 13,* 2010-2223.

Fischer-Kern, M., Buchheim, A., Hörz, S., Schuster, P., Doering, S., Kapusta, N., Taubner, S., Tmej, A., Rentrop, M., Buchheim, P., Fonagy, P. (2010). The relationship between personality organization, reflective functioning, and psychiatric classification in borderline personality disorder. *Psychoanalytic Psychology 27(4)*, 395-409.

Fischer-Kern M., Kapusta N.D., Doering S., Hörz S., Mikutta C., Aigner M. (2011). The relationship between personality organization and psychiatric classification in chronic pain patients. *Psychopathology*, *44,* 21-26.

Fromm-Reichmann, F. (1950). *Principles of intensive psychotherapy.* Chicago, IL: University of Chicago Press.

Hörz, S., (2007). *A prototype of borderline personality organization - Assessed by the Structured Interview of Personality Organization (STIPO).* Hamburg, Germany: Kovac.

Hörz, S., Stern, B., Caligor, E., Critchfield, K., Kernberg, O.F., Mertens, W., & Clarkin, J.F. (2009). A prototypical profile of borderline personality organization using the Structural Interview of Personality Organization (STIPO). *Journal of the American Psychoanalytic Association, 57(6),* 1464-1468.

Hörz, S., Rentrop, M., Fischer-Kern, M., Schuster, P., Kapusta, N., Buchheim, P., Doering, S., (2010). Strukturniveau und klinischer Schweregrad der Borderline Persönlichkeitsstörung [Personality structure and clinical severity of borderline personality disorder]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie, 56(2),* 136-149.

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.A. Levy, J.S. Ablon & H. Kächele (Eds.), *Psychodynamic psychotherapy research* (pp. 571-592). New York, NY: Humana.

Hörz-Sagstetter, S., Caligor, E., Preti, E., Stern, B., De Panfilis, C., & Clarkin, J.F. (2018). Clinician-guided assessment of personality using the Structural Interview and the Structured Interview of Personality Organization (STIPO). *Journal of Personality Assessment, 100(1),* 30-42.

Kernberg, O.F. (1981). Structural interviewing. *Psychiatric Clinics of North America, 4(1),* 169-195.

Kernberg, O.F. (1984). *Severe personality disorders: Psychotherapeutic strategies*. New Haven, CT: Yale University Press.

Kernberg OF, & Caligor E (2005). A psychoanalytic theory of personality disorders. In M.F. Lenzenweger & J.F. Clarkin (Eds.), *Major theories of personality disorder* (2nd ed., pp. 114-156). New York, NY: Guilford Press.

Leichsenring, F. (1997). Borderline-Personlichkeits-Inventar (BPI) Manual [Borderline Personality Inventory]. Gottingen, Germany: Hogrefe.

Lenzenweger, M.F., McClough, J.F., Clarkin, J.F., Kernberg, O.F. (2012). Exploring the interface of neurobehaviorally linked personality dimensions and personality organization in borderline personality disorder: The Multidimensional Personality Questionnaire and Inventory of Personality Organization. *Journal of Personality Disorders, 26(6),* 902-918.

Lingiardi, V., & McWilliams, N. (Eds.). (2017). *Psychodynamic Diagnostic Manual (2nd ed.).* New York, NY: Guilford Press.

Loranger, A. W. (1999). *International Personality Disorder Examination (IPDE).* Odessa, FL: Psychological Assessment Resources.

McKinnon, R., & Michels, R. (1971). *The psychiatric interview in clinical practice*. Philadelphia, PA: Saunders.

McKinnon, R., Michels, R., & Buckley, P. (2006). *The psychiatric interview in clinical practice* (2nd ed.). Washington, DC : American Psychiatric Publishing.

Powdermaker, F. (1948). The techniques of the initial interview and methods of teaching them. *American Journal of Psychiatry, 104(10),* 642-646.

Preti, E., Prunas, A., Sarno, I., & De Panfilis C. (2012). Proprietà psicometriche della STIPO [Psychometric properties of the STIPO]. In: F. Madeddu & E. Preti (Eds.), *La diagnosi strutturale di personalità secondo il modello di O.F. Kernberg. La versione italiana della Structured Interview of Personality Organization* (pp. 59-84). Milan, Italy: Raffaello Cortina.

Preti, E., Rottoli, C., Dainese, S., Di Pierro, R., Rancati, F., & Madeddu, F. (2015a). Personality structure features associated with early dropout in patients with substance-related disorders and comorbid personality disorders. *International Journal of Mental Health and Addiction, 13(4),* 536-547.

Preti, E., Prunas, A., De Panfilis, C., Marchesi, C., Madeddu, F., & Clarkin, J. F. (2015b). The facets of identity: Personality pathology assessment through the Inventory of Personality Organization. *Personality Disorders: Theory, Research, and Treatment, 6*(2), 129-140.

Preti, E., Di Pierro, R., Costantini, G., Benzi, I. M., De Panfilis, C., & Madeddu, F. (2018). Using the Structured Interview of Personality Organization for DSM–5 Level of Personality Functioning Rating performed by inexperienced Raters. *Journal of Personality Assessment, 100(6),* 621-629.

Rentrop, M., Zilker, T., Birkhofer, A., Lederle, A., Hörz, S. (2014). Psychiatric comorbidity and personality structure in patients with polyvalent addiction. *Psychopathology, 47(2),* 133-140

Shedler, J. & Westen, D., (2007). The Shedler-Westen Assessment Procedure (SWAP): Making personality diagnosis clinically meaningful. *Journal of Personality Assessment, 89(1),* 41-55.

Shedler, J. (2015). Integrating clinical and empirical perspectives on personality: The Shedler-Westen Assessment Procedure (SWAP). In: S.K. Huprich (Ed.), *Personality disorders: Toward theoretical and empirical integration in diagnosis and assessmen*t *(225-252).* Washington, DC: American Psychological Association.

Stern, B.L., Caligor, E., Clarkin, J.F., Critchfield, C., Hörz, 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* 2010; *92(1),* 35-44.

Sullivan, H. (1954). *The psychiatric interview.* New York, NY: Norton.

Verheul, R. (2006). Clinical utility of dimensional models for personality pathology. In: Widiger, T.A., Simonsen, E., Sirovatka, P.J., & Regier, D.A. (Eds.), *Dimensional models of personality disorders: Refining the research agenda for DSM-V (203-218)*. Washington, DC, American Psychiatric Association.

Verheul, R., Andrea, H., Berghout, C.C., Dolan, C., Busschbach, J.J.V., van der Kroft, P., Bateman, A.W., & Fonagy, P. (2008). Severity indices of personality problems (SIPP-118): Development, factor structure, reliability, and validity. *Psychological Assessment, 20(1),* 23-34.

Westen, D., & Muderrisoglu, S. (2003). Assessing personality disorders using a systematic clinical interview: Evaluation of an alternative to structured interviews. *Journal of Personality Disorders, 17(4),* 351-369.

Whitehorn, J. (1944). Guide to interviewing and clinical personality study. *Archives of Neurology and Psychiatry, 52(3),* 197-216.

Yeomans, F., Clarkin, J.F., Kernberg, O.F. (2015). *Transference-Focused psychotherapy for borderline personality disorder: A clinician's guide.* Washington, DC: American Psychiatric Publishing.

SECTION 2: STIPO-R INTERVIEW GUIDE

2.1 General Administration Issues

2.1.1 Periods of not being "one's normal self"

If, at the start of the interview, a person reports not having been his or her "normal self" for a significant period of time during the past five years, one must inquire about the nature of the disturbance. One generally wants to score the interview only for the time period in which the person was, in fact, his or her "normal self." This may mean that the period of time over which the respondent will be reporting will be less than the standard five years.

For our STIPO validation study we eliminated patients with lifetime diagnoses of schizophrenia and Bipolar I disorder. Our experience with the interview in the clinical samples we used, which initially included patients with these diagnoses, suggested that such patients, due to the influence of their symptoms, could not accurately reflect upon the experience or idea of their "normal self."

Many patients, particularly borderline patients, but also depressed and bipolar II patients, will answer this question affirmatively. This raises a question: is it that they are not their "normal selves" during parts of the five years, or that their "normal self" is in fact highly unstable, discontinuous, or otherwise disturbed? The key thing to discern at this point is the amount of time within the past five years was the respondent not "his or her normal self," and how his or her functioning differed from his or her normal functioning during that time. If your sense as the interviewer is that a discrete time can be identified in which the respondent's personality was discontinuous from what is normal for them, *due to interference from acute symptomatology or a severe traumatic event*, then one should exclude that discrete period from the five-year period. In the absence of a clear, discrete period of time in which the subject was not his or her "normal self" due to an acute symptomatic mental illness or a severe traumatic event, ask the subject to survey the entire five-year period.

The next question would be what is the minimum amount of "normal self" time upon which an interview can be reliably scored. For example, if a subject says that he or she was drug-addicted for two of the five years, would the remaining three years allow for a accurate reflection of the respondent's "normal self" or personality? At this point, we would say yes. For the sake of standardization in administration and norming, we currently recommend that there be at least a *three-year period* of "normal" functioning for the interviewer to assess; otherwise, we recommend that the interview be discontinued as it is unlikely that the data will reflect a true sense of respondent's personality.

2.1.2 A guide to mandatory question stems, follow-up probes, and optional probes

a. Mandatory questions. All mandatory STIPO questions are bolded in the interview. The interviewer *must* ask each of those item stems.

b. Probes. In several places in the interview you may see a prompt to Probe. This prompt will provide some general language to use, in your own words, to get further clarification as the primary item stem.

c. Additional language, non-bolded. Almost every question in the interview has additional text that can be added at the interviewer's discretion to obtain further clarification as to the primary item stem. Some of this additional language may be used, or none may be asked depending upon the answer to the main item stem. The ultimate goal of the interviewer is to be able to score the question, and the additional, non-bolded language is an attempt to standardize the language used by the interviewers in seeking further clarification of initial interview responses.

d. Conditional questions If yes,… If no,…..

Certain questions must be asked in follow up if either a yes or no response is given, most commonly "If yes, …..". All "If yes" and "if no" questions are mandatory questions.

e. Notes (Note: ……..)

Notes are to inform the interviewer about the essence of the question and to provide assistance with and structure for the interviewer in querying respondents about unclear answers. In some cases, the notes will contain general comments about the type of information that the interviewer should probe for, and in other cases specific language for the interviewers' questions is provided.

f. General, non-specified probes

In general, if the interviewer is uncertain about the scoring of a response, it is appropriate to ask any of the following probes:

> Does a recent example of this come to mind?
> Is this something you do frequently?
> Is this kind of behavior typical and frequent, or rare?

2.1.3 The 0-2 anchors

0 = pathology absent; the trait being queried is not present at all, or, if slightly present, has no impact on respondent's functioning.

1 = the trait being queried is present, and reflective of some pathology, sub-threshold; minor impairment

2 = the trait being assessed is clearly present and reflects significant to severe pathology; significant to severe impairment.

2 is really a broader category than 0 and 1 as it encompasses those who have the trait being assessed and who manifest some impairment as a result, and those whose lives are severely compromised by the trait.

Unless it is explicitly stated, the respondent is not expected to meet all of those features listed under a given anchor in order to score at that level. We simply provide a listing of the features that *could*, in various combinations, constitute a response at that level. When evaluating any of those characteristics and trying to distinguish, for example, a score of **1 from a **2**, the interviewer should consider the *frequency, intensity or severity, and pervasiveness* of the particular behavior or feeling being assessed.

2.1.4 How to deal with interdependent questions

For example, if the person has had no sexual relationships in the past 5 years, how does one score the question about "preoccupied with evaluating how much you get out of the relationship in relation to how much your partner gets out of it"? We concluded that we would score the item "9", which stands for "question skipped" or "not applicable" and that later in our computer scoring we would recode the responses into 2's, giving the person the lowest possible score. This scoring rule takes place in the following areas: work/school, friendships, romantic relationships, and sexual relationships.

2.1.5. Judgment calls and difficulty fitting a response to an anchor

From time to time a patient's response does not either speak to the question being asked, or neatly fit into the anchors provided. Under such circumstances, the interviewer should consider the following: "what is this person's response saying about the domain in question?" For example, for question #2, about the importance of studies / work, and the relation of studies/work to life goals; if the response does not answer the question precisely, or if it is not clear how the response fits the anchors, after asking each of the probes and following up as needed, then the interviewer should think "what is this response saying about the respondent's *capacity to invest?",* insofar as in this instance, the item falls under that sub-domain of Identity.

Second example. If, a respondent indicates that he or she has been moving from part-time job to part-time job, with no job lasting more than a few weeks or months, questions 1-3 make less sense than the overall notion that the person is simply not invested in work. In this case, one should again "*default to the domain*," scoring a 2 for each of the items 1-3, which reflects both that there really is no primary role, and, that the person's capacity to invest is severly compromised.

2.1.6 0-2 scales versus 1-5 overall ratings

There are two types of ratings that the interviewer is asked to make while scoring the STIPO. The 0-2 and the 1-5 systems are both linked closely to the content of the individual questions.

The 0-5 rating scales should reflect the interviewer's total clinical impression, based on all available information, verbal, non-verbal, and the respondent's level of superficiality across the questions in a given section. Furthermore, the interviewer can use the 5-point scale to weight questions differently; for example, the level of pathology expressed in one 2 rating may be

significantly more severe than that expressed in another response, also rated 2. Also, a subject may receive a **2** one only one question in a given section, but the interviewer may feel that this response alone reflects very significant pathology. This pathology would not be reflected as well in the 0-2 ratings, where all items are weighted equally in calculating the score; in contrast, the interviewer can weight that one item more heavily in determining calculating the level of pathology reflected in the 1-5 scale. The interviewer should make these ratings without any conscious effort to try and reconcile the 0-2 scores with the 1-5 ratings. The 1-5 rating should simply be made based on the interviewer's clinical judgment or sense of the respondent based on the questions in that section.

2.2 Guidelines for Specific STIPO Items

General comments:

- We use all available, observable information during the interview process (verbal and behavioral) to score a 0-2 rating. We do not make interpretations or inferences to help in scoring an item, *but if the respondent's verbal report is contradictory across the interview, we want to confront the discrepancy, and if the behavior contradicts a verbal report, we may choose to weight the behavior, which is also valid clinical information.*

- Even if one thinks that a second rater may not see it the same way, the task is to score it based on the actual information presented (verbal or otherwise).

- One should always revisit prior questions if new information contradicts an older reply.

Overview probe:

- Convey that it is an extended period of time, not just a couple of days.

- Criteria for exclusion is a marked divergence from normal self, with a clear decrement in one's typical functioning ("were you unable to function as your 'normal' self during this period"?)

- In response to "grandmother died", or some other focal stressor or trauma, assess the time period, and whether it was a marked divergence from normal self and how quickly the subject returned to baseline

Questions 1-3 Investment in Work / Studies, i.e., Primary Role:

- How to determine **primary role**:
  If the subject has had a primary role as **both a student and a worker** in the 5-year period, with substantial roles in both areas, then the interviewer is to **choose one or the other** based on:
  - Either the amount of time
  - Where the respondent would say their primary investment of time or energy was
  - If all else fails, just choose one.

- If the subject **has neither worked nor been in school** during the time period, when one would presume that either working or being in school should be their primary role, **score** all 9's for both sections.

  Example: A graduate student who did some part-time odd jobs, like 10 hours per week of temporary work, or worked in a student recreation center or retail shop during graduate school – only score *school* for that person, with the rationale being that *school wa*s clearly

the more significant primary role during that time. If, however, the person was a full-time graduate student while also working in a 20-hour per week research or teaching assistantship, that is also a significant work-role during that time, so one would need to ask which role the respondent felt was primary, or have the respondent or interviewer arbitrarily choose.

- If the subject is only **working part-time** (e.g., 10 hours per week) yet says that he/she is "very effective, very ambitious":
    - The general impression might be that they are not as invested as they say, or perhaps not capable of investing, which is the main point of the section. Ideally, probing more on the individual items will reveal whether this is the case, which would result in a 1, at most, and more likely a 2.
    - Probe as to why they not working more. Question whether if they were working more they might be as effective (e.g., "too much stress", "I need my time", etc.) or whether they are just are "very ambitious" but only in a 10 hour per week way, i.e., which is NOT that ambitious
    - For sure, this warrants a *lower* score in the 5-point rating

- If respondent is doing **family work**, one can ask the same question. Probe effectiveness, ambition (is this chosen, are there "goals"), and is there satisfaction (the **score** "0" here would be the realistic one, e.g., "there are moments when I love it and moments when I want to kill myself out of boredom or frustration." This would be normal)

- **Extenuating circumstances.** It is important to query participants who are not currently working as to why. For some, there may be legitimate, extenuating circumstances that have prevented them from working for a period within the 5 years, for example, a severe economic downturn, or an injury that removed them from the workplace. In such cases, the individual should not necessarily be penalized for the time *not* working in the 5 year period. If the best assessment is that the absence from work is unintentional, in circumstances like this the capacity to invest would be best assessed using only the time within the 5 years on which the participant *was* working, assuming there is a reasonable sample of a year or more over which to evaluate effectiveness, consistency, and pleasure.
-

Questions 1 Effectiveness:

- If a person received good external reviews for work effectiveness but still says they are working below their potential, one should separate out their poor self-esteem and self- criticism from the actual, objective performance issue. So, we would still **score** this a "0" if all the other measures of effectiveness check out, with exception of their self-evaluation.

Question 2 Ambition / Goals:

- If responded says something along the lines of "I want to put my kids through college," or "I just want to make money to support my family," this still **scores** a "0"as it reflects a capacity to invest in something for a goal, which indeed reflects ambition. Shifts within a field, e.g., from professor of finance to financial consulting can be **scored** a "0" or "1".

Question 4 Recreation:

- What the interviewer is looking for here is not just something the subject does, but *something they do and learn about*, spend time thinking about, and engaging in it *even when they're not doing it*. Let's say they cook; to score a 0, we'd ask if they are researching food and cooking

blogs, magazines, etc.? Just because it is something that they do, even if they do it regularly and consistently over time, that may not be sufficient for it to be a bonafide, deep, recreational investment. At best, such interests would be **scored** a "1".

- **Yoga, pilates, gym, reading, movies**: Not in itself **scored** a "0", even if it is consistent over time and takes up considerable, regular time. If yoga, pilates, gym, running, for example, are just to stay in shape, that's not a recreational investment that qualifies for a 0. If the person says that they take an avid interest in the gym, such that they read about fitness, make efforts regularly to learn about fitness and healthy living, i.e., that it is in a way more than just working out for one's health but a broader interest, then this is moving more into the lines of a fitness / health interest, i.e., a **score** of "0". If they read or watch movies for pleasure because they "like the movies", that's something they do, not something they're necessarily invested in; that also is *not* a recreational investment. If, however, they read about literature, follow specific writers, belong to book / reading groups, i.e., something more than "I just like to read", this moves towards a 0. If they just are avid readers, this would be **scored** a "1".

- **AA:** People sometimes talk about their involvement in Alcoholics Anonymous. This could either be seen as treatment, which is not a recreational interest, or, depending on how involved the person is in AA, a more serious recreational investment. For example, a person who has been sober for many years, who attends several meetings a week, not out of necessity but because of his or her commitment to AA, who takes on leadership roles and sponsors other alcoholics – this is clearly moving in the direction of a serious investment, i.e., a **score** of "0".

- If no sustained interest is identified, one might follow up asking: "Tell me about your free time; do you enjoy it and find your recreational time fulfilling, or is there is a lot of unstructured free time that you find yourself not knowing what to do, and feeling unsatisfactory or unenjoyable to you?" If the respondent enjoys his or her free time, keeps oneself occupied socially or recreationally but without a sustained interest as defined above, but without significant boredom or distress, that can also be scored a 0.

- If a person does not have clear activities, but feels somewhat engaged and experience some pleasure in free / recreational time, **score** this a "1". If a person has significant unstructured free / recreational time from which they derive little to no satisfaction **score** this a "2".

Questions 5/6 Sense of Self, Coherence and Continuity/ Ambivalence; 12/13 Representation of Other, Superficiality vs. Depth/ Ambivalence:

- Open ended representation of self-probe. Items # 5/6 and 15/16 are scored for three qualities, reflecting representations that are:
    - Superficial vs. Deep / Nuanced. Descriptions consisting only of superlatives, e.g., "the most amazing", "so wonderful", points to some defensive distortion or idealization. One should ask for greater elaboration (elicit example per the probes), but if no greater depth is elicited, this **scores** a 2.
    - Realistic / Integrated vs. Distorted / Polarized (idealized / devalued), whether the person can identify both positive and negative qualities in more than a caricatured manner. This does not have to be the deepest, most nuanced description of each. What one is after here is the ability to think about and describe both positive and negative in a manner that is somewhat realistic and elaborated. This alone **scores** a 0

for items 6 and 13.. Descriptions with less depth, it **score** a 1. Inability to think of either positive or negative **scores** a 2.

- Do not deviate from the script. Ask the first question ("Tell me about yourself as a person….") and then wait for the respondent's full reply. Then move onto the second probe (Is there anything else you can…..), and ask it exactly as it is listed.

- Next, the interviewer needs to probe for depth and elaboration. If the respondent simply lists qualities, the interviewer can probe or depth with probe #1 ("you've used several adjectives to describe yourself.…). Similarly, if one or two qualities were discussed but not in sufficient depth, the interviewer can respond with probe #2 (inquire about one or more adjectives…). In either case, as if the respondent can "fill in the description a bit, perhaps bringing it to life with an example or story that illustrates that quality."
-
- This should be done (eliciting example or story) for one positive and one negative quality.

- Ratings for both 5 and 6 are made after ALL probes are given.

- Beware of false "negative" qualities, e.g.., "I'm too modest", "I'm very self-critical", which could be disguised expression of grandiosity or underscoring positive qualities without any owning of something negative

- If deciding between a **score** of "0" and a "1", the ease with which one can bring an example would tilt to scoring to a "0", whereas a poverty in the narrative or difficulty bringing an example would tilt to a "1".

Question 7 Consistent Sense of Self in Presence:

- If respondent behaves in an unpredictable/ erratic manner during interview this can either be inquired about or taken into consideration by the interview in scoring the item, even if it contradicts the verbal report.

- The respondent may report playing different roles as called for by specific situations, e.g., needing to be aggressive in a business or legal setting, versus being more accommodating and sensitive in one's intimate relationships – but does not report feeling like a different person across those situations. The key to this item is indeed whether they begin to feel like a different person or try to take on a different personality or self in a given situation, such that their sense of self and who they are actually is different in those situations.

- Some people will say that they are "predictably unpredictable." That is the definition of what we're looking for in this item and would generally qualify as a **score** of "2". If the person says "No, I'm not seen as erratic, my friends know that I'm all over the place, they expect me to act erratic", that's either a **score** of a "1" or "2" depending on how erratic, severe, or pervasive this quality is.

- It is normal for people to present somewhat differently across work and personal lives. The question would be whether they present differently *within* those settings, and whether they feel like a different person across those settings. Feeling like a different person across those settings, and/or presenting variably within each setting, suggests identity diffusion, i.e. a **score** of "2".

Question 8 Self-Tastes/ Opinions:

- Open versus impressionable. Do you take on the opinions of others as if they were your own, not really having a sense of what tastes and preferences come from inside yourself. The key is whether you look to the outside because you have no sense at all inside. If opinions are unformed this does not necessarily **score** as a "2"; it would **score** as a "2", however, if the person consistently looked to the outside to get a sense of those opinions and then took them as his or her own. The key here is judging whether the person needs to consistently look to the outside to get a sense of what he or she feels inside, or whether most of the person's opinions, tastes, preferences are internally derived. Again, it's totally fine if a person says that they don't have strong opinions – what the interviewer is looking for is how comfortable they are with that, and whether they take on as their own the opinions of others to cover up for their lack of opinions or tastes.

Question 9 Narcissistic Supplies:

- If the respondent indicates that they feel empty or down when they are not getting attention/admiration, one may follow up by asking: "When you get deflated like that, how long can that last, and how badly does it feel?"

Question 10 Self in Intimate Relationships:

- As being "flexible" can be adaptive, accommodating one's partner still can **score** a "0". Regular to exclusive submission to one's partner to avoid conflict, even if stated as "being flexible", moves **scores** to a "1" or "2" depending on severity and pervasiveness. Regular submission to one's partners preferences / needs **scores** a "1", almost exclusive submission to one's partner's preferences / needs **scores** a "2".

Question 11 Self-Esteem:

- This is about the degree of fluctuation, the sense of stability in the sense of self-esteem (not valence but stability).

Question 15 Others' Feelings about the Self:

- This is about cognitive confusion related to the difficulty a person experiences in assessing how others view him or her. One thing we tend to pick up here is the subject's projected self-criticism, which is different.

Questions 16-18 Friendships:

- If there are no friends at all, or the friends described result in a **score** of "2" on #16, then **score** both #17 and #18 "2".

Question 16 Friendships Presence:

- As the item is simply a measure of social connectedness versus isolation one may wish to ask additionally if respondent feels socially isolated or connected.

- If deciding between a **score** of "1" and "2", take participants feelings of social connectedness versus isolation into account.

Question 17 Friendships Closeness:

- Probe the question for the two people closest to the respondent.

Question 18 Friendships Temporal Stability:

- If having difficulty discerning between a **score** of "1" and "2", one may follow up with: "Why is it that your friend group over time has shifted in this way?"

Questions 20 and 21 Intimate Relations:

- If the respondent has had no significant romantic relationships in the past 5 years (#20), then skip question 21. Similarly, if question #20 is **scored** a "2", meaning the romantic "relationship" was significantly brief or flawed, then we **score** #21 also as a "2".

- "Sexual" is not limited to intercourse.

- In determining a "significant" relationship in terms of duration, to allow for standardization across interviewers, we are proposing a relationship lasting 8 months or longer.

Question 22 Sexual Activity:

- By satisfaction we typically mean some combination of frequency, pleasure, comfort and connection.

- Some theoretical notes: Normal = sexual love + sexual pleasure; Neurotic = less frequent sexual activity, inhibition in pleasure, difficultly combining love and sex; Borderline = polymorphously perverse, sex in service of aggression. What is intended here is attempting to strike a balance between several issues: Is the person having sex, is the sex in the context of an ongoing relationship, and is the subject satisfied by and able to enjoy the sexual experience. If any of these aspects are seriously flawed, the response should lean towards a **score** of "2".

Question 23 Shyness about Sex:

- One may follow up with: "Would you consider yourself inhibited?"

- If not sexually active (#22) at all in the past 5 years, then skip question 23.

Question 26 Boredom:

- What the interviewer is after here is the subject's ability to sustain relationships over time. Losing interest relates to patterns of idealization / devaluation, as opposed to either growing apart, or realizing over time that the relationship was not working or meant to be.

- Similarly, some respondents report feeling "disgusted" by their partners over time. Although not related to boredom per se, this response still reflects a difficulty with the capacity for an internal investment in the other, and thus one would **score** this a "2".

Question 29 Economic View of Relationships:

- Here one is looking for respondents who either take a pervasively exploitive attitude or approach in their object relations, always needing to be getting the most out of their relationship partners in relation to what they are giving themselves, i.e., the exploiter attitude, OR, the respondent who is pervasively masochistic, always in a giving position, i.e., preoccupied with getting less out of the relationship.

- If the respondent is always insistent on the partner getting more, *or of submitting oneself,* **score** this a "2".

Question 31 Paranoia:

- The distinction to draw between a **score** of "1" and a "2" might be shame over aspects of self they do not want to reveal ("1"), versus the need to guard versus manipulation, which is a more narcissistic / borderline tendency ("2").

- One may wish to clarify by: "I'm speaking here not just about your close relationships, but about your general approach to people and the world; would you characterize yourself as more open, or closed off?"

Question 33 Black and White Thinking:

- One may follow up with: "Are you the kind of person who can easily see both sides of an issue, the nuance in things?"

Question 35 Idealization / Devaluation II:

- One may follow up with: "Some people tend to be "prickly" or hypersensitive or reactive, and others tend to be more easygoing; which is more characteristic of you?"

Question 37 Anticipation / Planning:

- If the respondent is not proactive with planning, one may follow up with: "Is your life generally well organized and effective, specifically, are you losing things, always running late, missing deadlines, etc.?"

- A respondent who does not endorse proactive planning as described, but copes effectively still **scores** a "0".

- Some of our study participants do not have lives that involve significant commitments or expectations, i.e., there is not much to be organized or proactive about (no job, no school, no responsibilities for child-care, etc.). Thus, they say that they don't engage in proactive coping, but do not experience stress or because of this. These participants should still score *"2" due to the presumption that having a life devoid of stress and commitment reflects significant rigidity and poor coping.*

Question 38 Suppression:

- One may follow up with: "Some people call this compartmentalizing, where you can put troubling things away long enough to go on and get stuff done; is that something you can do?"

Question 40 Perfectionism:

- If a person is not conscientious at all or there are no circumstances in life where such ambition can reasonably be applied, **score** this a "2".

Question 41 Aggression Self-Neglect:

- "I don't get enough sleep," "I don't eat as well as I should" **score** as "0".

Question 44 Suicidality:

- If a person indicates that they have not made any sucide attempts in the last year, one may follow up by asking: "Do you frequently fantasize about suicide?"

- A significant preoccupation with suicidality **scores** a "1", a significant to severe preoccupation with suicidality **scores** a "2".

Question 46 Envy:

- One may also ask whether the respondent feels resentful towards others who succeed or accomplish something.

Question 50 Moral Action:

- The interviewer is attempting to assess the presence of an internalized, yet flexible, moral code, internal to that respondent. Example: A religious respondent who says it is wrong to have pre-marital sex, but engages in some sexual activity, would still **score** a "0", to the extent to which he or she is engaged and wrestling with an internal moral code.

Question 51 Internalized Moral Values:

- Individuals governed by "religious" codes of moral conduct may still **score** a "1" if that code is governed by a fear of punishment more than an *articulated* sense of right and wrong.

Question 52 Guilt:

- Guilt means, "I have wronged / hurt someone else and I feel badly for how I have hurt them with emphasis on concern for the other." NOT: "I feel badly b/c I let myself down." VERSUS - "I feel guilty" but meaning, I know I behaved poorly and am focused on how bad I am, or how wrong I was and feel a need to punish or to avoid myself, which is a paranoid / shame dynamic. One is looking for a capacity for guilt in the depressive position sense: I have hurt others, and am regretful and concerned about damage done to them; there should be reflection/effort on doing things differently in the future or on making amends.

- Needs a clear examples to get a **score** of "1"; where there is some concern that the behavior has adversely impacted others.

- The example may actually be less guilt than self-criticism, e.g., "I don't go gym", "I should have been a better mother when my kids were young." This does not **score** a "0" but a "1".

- If subject says "Others have said that i'm guilty of x": follow up with "well how do *you* feel about that?"

- If the respondent does not provide an expample at all we **score** this a "9".