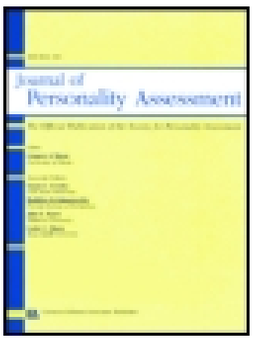

Routledge
Taylor & Francis Group

Journal of Personality Assessment

ISSN: 0022-3891 (Print) 1532-7752 (Online) Journal homepage: http://www.tandfonline.com/loi/hjpa20

Relationship of Pretreatment Rorschach Factors to Symptoms, Quality of Life, and Real-Life Functioning in a 3-Year Follow-Up of Traumatized Refugee Patients

Marianne Opaas, Ellen Hartmann, Tore Wentzel-Larsen & Sverre Varvin

To cite this article: Marianne Opaas, Ellen Hartmann, Tore Wentzel-Larsen & Sverre Varvin (2015): Relationship of Pretreatment Rorschach Factors to Symptoms, Quality of Life, and Real-Life Functioning in a 3-Year Follow-Up of Traumatized Refugee Patients, Journal of Personality Assessment, DOI: 10.1080/00223891.2015.1089247

To link to this article: http://dx.doi.org/10.1080/00223891.2015.1089247

Published online: 03 Nov 2015.

JOURNAL OF PERSONALITY ASSESSMENT
http://dx.doi.org/10.1080/00223891.2015.1089247

Routledge
Taylor & Francis Group

Relationship of Pretreatment Rorschach Factors to Symptoms, Quality of Life, and Real-Life Functioning in a 3-Year Follow-Up of Traumatized Refugee Patients

Marianne Opaas,[1] Ellen Hartmann,[2] Tore Wentzel-Larsen,[1,3] and Sverre Varvin[4]

[1]Norwegian Centre for Violence and Traumatic Stress Studies, Oslo, Norway; [2]Department of Psychology, University of Oslo, Norway; [3]Center for Child and Adolescent Mental Health, Eastern and Southern Norway, Oslo, Norway; [4]Department of Nursing and Health Promotion, Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences, Oslo, Norway

ABSTRACT

Response to mental health treatment varies highly among refugee patients. Research has not established which factors relate to differences in outcome. This study is a follow-up of Opaas and Hartmann's (2013) Rorschach Inkblot Method (RIM; Exner, 2003) pretreatment study of traumatized refugees, where 2 RIM principal components, Trauma Response and Reality Testing, were found descriptive of participants' trauma-related personality functioning. This study's aims were to examine relationships of the RIM components with measures of anxiety, depression, posttraumatic stress, quality of life (QOL), employment, and exile language skills throughout 3 years. We found that impaired Reality Testing was related to more mental health symptoms and poorer QOL; furthermore, individuals with adequate Reality Testing improved in posttraumatic stress symptoms the first year and retained their improvement. Individuals with impaired Reality Testing deteriorated the first year and improved only slightly the next 2 years. The results of this study imply that traumatized refugee patients with impaired Reality Testing might need specific treatment approaches. Research follow-up periods should be long enough to detect changes. The reality testing impairment revealed by the RIM, mainly perceptual in quality, might not be easily detected by diagnostic interviews and self-report.

ARTICLE HISTORY
Received 19 May 2014
Revised 8 July 2015

Exposure to torture, threats to one's life, and the additive effects of other adverse experiences of war and human rights violations (HRVs) have repeatedly been related to a heightened risk of mental health disorder (e.g., Steel, Silove, Phan, & Bauman, 2002). Many traumatized refugees remain with severe symptoms and poor functioning long after the traumatizing events and arrival in the exile country (e.g., Vaage et al., 2010).

Refugees often have a history of repeated wars or other adverse conditions. Intergenerational trauma associated with decades or centuries of oppression and wars can upset child care and parent–child attachment (e.g., Van Ee, Kleber, & Mooren, 2012). Severe or repeated adverse experiences in vital periods of development can seriously affect emotional and physiological regulation and interpersonal relating, and lead to increased vulnerability to later stressors (e.g., Fonagy & Target, 2005; Hinton & Kirmayer, 2013; Schore, 2002). Previous adverse experiences might thus aggravate the effects of war experiences in adult years.

The proposed diagnoses "*complex posttraumatic stress disorder (C–PTSD)*" or "*complex trauma*" (Curtois, 2004; Herman, 1992) represent efforts to incorporate the disturbances in intra- and interpersonal functioning, the difficulties in overall adjustment, and the comorbid disorders often found in individuals exposed to early, severe, or repeated trauma (Schottenbauer, Glass, Arnkoff, & Gray, 2008). Further, the ICD–10 diagnosis "*enduring personality change after catastrophic experience*" (F62.0; World Health Organisation Geneva, 1994) accounts for lasting, serious alterations in personality functioning after traumatic exposure. Hinton and Kirmayer (2013) recommended focusing on negative affective states rather than PTSD per se, arguing that trauma commonly results in a wide range of negative states as well as PTSD.

Treatment of traumatized patients

Psychotherapies of various orientations have been found effective for PTSD, but severity of the condition and other personal characteristics affect the rate of change (e.g., Lambert, 2013). Positive outcome can be low in patient groups with complex and chronic posttraumatic stress, comorbidity, and more diverse constellations of symptoms. Treatment dropout might be high. Lambert (2013) estimated around 30% dropout in this patient group, compared to around 20% in the general patient population (Swift & Greenberg, 2012). Treatment outcome studies of traumatized refugees and asylum seekers are scarce, but indications of positive treatment gains are found in systematic overviews of randomized, controlled studies (e.g., Crumlish & O'Rourke, 2010), naturalistic studies (e.g., van Wyk & Schweitzer, 2014), and in an overview covering 30 years of intervention studies (McFarlane & Kaplan, 2012). However,

CONTACT Marianne Opaas marianne.opaas@nkvts.no Norwegian Centre for Violence and Traumatic Stress Studies (NKVTS), Postboks 181 Nydalen, 0409 Oslo, Norway.

when studying individual gains as well as group-level improvement, the rate of nonresponse was found to be high. Neither past trauma, sociodemographics, nor differences in treatment significantly explained these differences in outcome (Boehnlein et al., 2004; Carlsson, Olsen, Kastrup, & Mortensen, 2010).

Predictors of treatment outcome

Psychotherapy research points to comorbidity, initial functional impairment, and duration of the current episode as predictors of poorer outcome (Clarkin & Levy, 2004), which might explain poorer treatment gains in traumatized refugees with comorbid conditions. Complex interactions are at play in psychotherapy, between the personal characteristics of the patient and the therapist, their match, and the therapeutic approach. Still, patient personal characteristics are more important to outcome than therapist characteristics and treatment method (Bohart & Wade, 2013; Norcross & Lambert, 2011). Patient qualities, such as secure attachment and high ego strength, affect outcome positively but are also connected with healthier functioning overall. Personality-related variables, such as patient's coping style and openness to emotion and inner experiences, seem to interact with the personal style of the therapist and the treatment approach to yield different results (Bohart & Wade, 2013). Refugees' unfamiliarity with psychotherapy can make relationships even more complex.

The Rorschach Inkblot Method

The Rorschach Inkblot Method (RIM; Exner, 2003) is a performance-based personality assessment method, designed to assess a combination of implicit and explicit processes, reflecting underlying psychological characteristics. RIM requires the participant to use perceptual, cognitive, and affective resources to solve the unfamiliar task of proposing what 10 cards with suggestive, but nonfigurative inkblots might look like (Meyer, Viglione, Mihura, Erard, & Erdberg, 2011; Weiner, 2003). The accuracy of RIM results does not depend on a person's self-knowledge, as self-report measures do, and is less vulnerable to intentional distortion of responses (Hartmann & Hartmann, 2014). The RIM method can produce significant data that are not accessible through more direct and less time-consuming instruments (Meyer & Viglione, 2008). RIM results correspond better with more objective, functional measures, like observer ratings and measures of perceptual and cognitive disturbances, than with introspectively assessed criteria, such as self-report on questionnaires (e.g., Bornstein, 2012; Mihura, Meyer, Dumitrascu, & Bombel, 2013). Grønnerød's (2003) examination of RIM variables showed high test–retest stability, although stability was higher for variables reflecting trait-like aspects of personality, and lower for variables reflecting state-like features of personality. Comprehensive research has documented the reliability and validity of RIM variables (e.g., Mihura et al., 2013), as well as their cross-cultural applicability (Shaffer, Erdberg, & Meyer, 2007). Its composition of visual, nonverbal stimuli with no need for translation of individual items makes it especially suitable as a cross-cultural instrument (e.g., Allen & Dana, 2004).

In traumatized individuals, RIM stimuli tend to elicit trauma-related images and feelings (Kaser-Boyd & Evans, 2008), which enable communication of traumatic experiences in an indirect way, allowing observation of the impact of traumatic recollections on cognition, affect, and behavior (Levin & Reis, 1997). Studies of traumatized individuals have identified RIM indications of a biphasic response to trauma (van der Kolk, 1994; van der Kolk & Ducey, 1989), in which traumatic intrusions, hyperarousal, and emotional flooding alternate with emotional and cognitive avoidance, numbing, and constriction. Ephraim (2002) and Evans (2008) found indications of flooding or constriction shifting from one RIM response to the next, whereas Opaas and Hartmann's (2013) study found that either flooded or constricted characteristics tended to dominate the RIM protocols of the participants at the time of assessment. Viglione, Towns, and Lindshield (2012) observed how RIM responses reveal the struggle between loss of control and overcontrol, which they considered crucial to posttraumatic reactions.

Several authors have pointed to the value of the RIM as a tool for investigating how traumatization might affect a range of psychological functions (e.g., Ephraim, 2002; Kaser-Boyd & Evans, 2008). Further, brief protocols ($R < 14$) have been found to be characteristic of patients with PTSD (e.g., Arnon, Maoz, Gazit, & Klein, 2011), and to still be "interpretatively robust," meaning that they do not possess the high Lambda (indicating a tendency to simplify complex stimuli fields) and low Blends (indicating little richness and flexibility in thinking) that are often found in brief records. Brief responses have been interpreted as the patients' attempt to limit and escape painful associations (Brand, Armstrong, & Loewenstein, 2006). Several authors have dissuaded exclusion of short protocols in trauma populations as this could bias the findings (e.g., Arnon et al., 2011).

The RIM has been associated with behavioral and relational outcomes that develop over time (Viglione, 1999). Grønnerød's (2004) meta-analysis revealed that the RIM was a valid indicator of personality changes following psychotherapy and that RIM variables related to stress, control capacity, and affect changed the most in the course of therapy. Effect sizes of differences between test and retest RIM scores increased with therapy intensity and duration. Several variables and indexes of the RIM indicate psychologically healthy functioning; for example, the Ego Impairment Index (EII; Viglione, Perry, Giromini, & Meyer, 2011). Furthermore, Graceffo, Mihura, and Meyer (2014) found the Mutuality of Autonomy Scale (MA/MAO; Urist 1977) to be a valid measure of relational health as well as of general psychopathology across clinical and nonclinical samples. There is, however, limited research on how RIM variables might predict change. The Rorschach Prognostic Rating Scale (RPRS; Klopfer, Kirkner, Wisham, & Baker, 1951), the EII, the MOA, and Holt's (1977) Primary Process Aggression Score (AGG1) have predicted treatment outcome in psychiatric adult and child populations (e.g., Meyer & Handler, 1997; Perry & Viglione, 1991; Smith, Van Ryzin, Fowler, & Handler, 2014; Stokes et al., 2003). Hilsenroth, Handler, Toman, and Padawer (1995) found the RIM Comprehensive System (CS; Exner, 2003) variables Cooperative Movement (COP), Texture-Shading (T), and Aggressive Movement (AG) to discriminate between patients who terminated psychotherapy prematurely and those who stayed on in treatment. Fowler et al. (2004) found that changes in A1 (AGG1) and MOA measures at retest

correlated significantly with behavioral changes according to the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed. [*DSM−IV*]; American Psychiatric Association, 1994) Axis V scales, concluding that this supported the ecological validity of the changes in Rorschach scores. Apparently, no single Rorschach measure or index has been widely used to evaluate treatment outcome or general improvement.

The index study

This study is a follow-up of Opaas and Hartmann's (2013; called *the index study* henceforth) factor-analytic RIM study of 51 traumatized patients (64.7% male) with a refugee background. Participants were recruited during the years 2006 to 2009 at specialist mental health outpatient services with public funding. The inclusion criteria were adult patients (age 18−65) with refugee background and mental health disorder at least partly related to traumatic experiences of war, persecution, and flight. A formal PTSD diagnosis was not required. Present and severe psychosis or drug problems were exclusion criteria.

Seventy-two refugee patients were evaluated as eligible, 75.0% ($n = 54$) consented to take part in the study, and 70.8% ($n = 51$) met the further requirement of a minimum of seven responses to the RIM. Mean age was 39.4 ($SD = 8.0$), 66.7% were married, and 82.4% had children. They came from 15 different countries in Asia, Eastern Europe, and Africa (56.9% from the Middle East). Their mean stay in Norway was 11.2 years ($SD = 6.3$), most had Norwegian citizenship, mean education in their country of origin was 9.7 years ($SD = 4.4$), and few had completed any further education in Norway.

The first and last authors, experienced clinicians external to the treatment sites, assessed the participants at treatment start with the RIM, qualitative interviews, and with questionnaires about mental health symptoms and quality of life. We used professional interpreters when needed. The participants reported on average 16.7 ($SD = 6.4$) different kinds of potentially traumatic experiences of war or HRVs. Most had experienced military attacks, being forced to evacuate, and had experienced the violent death, murder, kidnapping, or disappearance of family or friends; 74.5% had been close to death and dying, 52.9% had been tortured, and 22.0% had been raped. At treatment start, 81.6% qualified for PTSD, 96.0% had clinically significant anxiety, and 98.0% qualified for major depressive disorder. The participants' quality of life (QOL) was low, with scores related to their experience of psychological health, physical health, and social relationships in the lower third of the scale.

The first author conducted the RIM in the third or fourth assessment session to allow time for some trust to develop. Administration followed the CS guidelines, moderated by more prompting and inclusion of new responses during inquiry, due to the few responses given by many participants. Thirty-three percent of the RIM protocols were still brief (7−13 responses), but without evidence of noncollaboration. The first and second author jointly scored all the protocols. An external RIM expert independently scored 20 randomly drawn protocols. Intraclass correlation scores ranged from .79 to .98. RIM variables for the principal component analysis (PCA) were selected by their established (e.g., Kaser-Boyd & Evans, 2008) and potential relevance to traumatization, by their interpretational validity (e.g., Mihura et al., 2013), and by their parametric qualities for statistical analyses (e.g., Meyer, Viglione, & Exner, 2001), and were all controlled for R.

The PCA resulted in two strong components, assumed to indicate underlying personality dimensions, accounting for 59.7% of the variance. The first component, *Trauma Response*, explained 40.8% of the variance; the second component, *Reality Testing*, explained 18.8% of the variance. The Trauma Response component consisted of the RIM variables Blends, CF+C, m, M, the Trauma Content Index (TCI), and Severe Cognitive Codes (SevCog), all positively weighted, and F%, negatively weighted. The Reality Testing component consisted of the variables FQo%, D, and R, all positively weighted, and FQ−%, negatively weighted. Blends (an indication of complexity), CF+C (poorly modulated emotions), m (fear of forces outside personal control), M (resourceful interest in or vigilance toward people), F% (a tendency to simplify or not notice complex external or internal information), D (attention to obvious parts of the environment), and R (verbal productivity, task engagement, and perceptual and associative capacity) were all derived from the CS. The TCI (sum of images of blood, anatomy, sex, aggressive action, and morbid responses, divided by R; indicating intrusive imagery) was derived from Armstrong and Loewenstein (1990). SevCog (suggestive of severe thought disturbances), FQo% (reality testing and conventionality of perception), and FQ−% (distorted perception) were derived from the Rorschach Performance Assessment System (R−PAS; Meyer et al., 2011).

The component scores, a standardized measure based on the relative contribution of each RIM variable, ranged from −1.82 to 2.60 on Trauma Response and from −2.50 to 1.52 on Reality Testing. Positive values of Trauma Response, *flooding,* represented RIM indications of being cognitively and emotionally overwhelmed by traumatic intrusions, trauma-related thoughts, memories, and strong emotions. Negative values, *constriction,* represented RIM indications of restricted emotional, cognitive, associative, and verbal activity. Reality Testing was termed *adequate* when values were positive, indicating an ability to accurately perceive and recognize obvious features of the environment, and *impaired* when negative, indicating less adequate perception, limited ability to recognize obvious aspects of the perceptual field, and a tendency to give few responses. Participants with a flooded Trauma Response on the RIM had significantly more reexperiencing symptoms of PTSD, whereas participants with impaired Reality Testing had more anxiety and poorer QOL in physical health, psychological health, and social relationships. For details, see Opaas and Hartmann (2013). Choices and limitations of the index study constitute the terms of this work.

The follow-up study

Aims and hypotheses

This study comprises a 1-year and 3-year follow-up of the traumatized refugee patients of the index study. We aimed

at further studying the characteristics of the RIM components of the index study and their relationships to mental health, QOL, and functional outcome over the 3 years. We expected that these components, Trauma Response and Reality Testing, describing two trauma-related dimensions of the participants' personality functioning, would be related to their mental health and QOL during follow-up, with how they responded to treatment, and with their course of change. Findings might contribute to better understanding of personal characteristics associated with mental health, response to treatment, and the capacity to improve under the influence also of any favorable or adverse factors outside therapy. Moreover, findings could generate hypotheses for further studies and inform treatment planning for similar patients.

First, we more closely evaluated the RIM components and the relationship between them. We understood the Reality Testing component as linear and one-directional, with increasingly positive scores representing increasingly more adequate reality testing (i.e., good functioning). Our first hypothesis was, therefore, that high and positive Reality Testing scores corresponded with normal and healthy personality functioning, whereas lower values increasingly indicated more problematic personality functioning. Neither direction of the Trauma Response component, however, seemed to represent good functioning; high negative values indicated more severe constriction and high positive values indicated more severe flooding. Our next hypothesis was, therefore, that midrange Trauma Response scores represented more normal or healthy personality functioning, whereas more positive (flooding) or more negative (constriction) scores represented less normal or less healthy personality functioning. Further, even though the components were relatively independent of each other in the PCA of the index study (also when using oblique rotation; $r = -.11$), we wanted to test the hypothesis that more "extreme" values of Trauma Response, in either direction, related to poorer Reality Testing.

Second, we aimed at investigating the relationships of the pretreatment RIM components, separately and in interaction, with symptom severity of anxiety, depression, and posttraumatic stress, level of subjectively experienced QOL, employment, and Norwegian communication skills, and with the change in these outcome variables during the 3 years of follow-up. Our hypothesis was that participants with more impaired Reality Testing would have more mental health symptoms and poorer QOL and be less likely to be employed or be able to communicate with us in Norwegian during follow-up, compared to participants with adequate Reality Testing; and similarly for participants with more constricted or more flooded Trauma Response, compared to participants with midrange Trauma Response, scores. We did not have any hypotheses regarding how the participants' scores on the components would relate to the rate or amount of change.

Our third aim was to investigate the relationships of the RIM components with frequency of therapy sessions and duration of treatment. From interviews with the therapists, we knew that they mostly scheduled weekly or biweekly sessions with their patients, but many patients' attendance was, for various reasons, less frequent.

Method

Participants

Participants were the same as in the index study. At treatment start (T1) there were 51 participants; at 1-year follow-up (T2), 50 participants (the one missing participant came for the next follow-up); and at the 3-year follow-up (T3), 46 participants (90.2% of the original sample). Even if participants dropped out of or terminated their treatment early, they mostly did not drop out of our study. Among the five that did not meet at T3 assessment, we knew (through telephone contact with them or through other sources) that one was working at T3 and three were not. Their situation seemed about the same as when they met with us at T2. We had no updated information about the fifth participant.

Treatment

All patients were offered psychotherapy, which to a certain extent addressed the patient's life history, traumatic experiences, current relationship issues, and problems in daily living. The therapies were of multiple theoretical orientations, not manualized, and not part of any special program for refugees. We were not able to record medication systematically. Most had used psychopharmacological medication at some time, but reported negative side effects and few positive effects of medication. Nonadherence to prescriptions was typical, with overuse of sleeping medication and discontinuation of other medication.

Psychotherapy was performed by specialists in clinical psychology (58.8%, $n = 30$), specialists in psychiatry (25.5%, $n = 13$), and by licensed clinical psychologists, medical doctors in training for the psychiatric specialty, and nurses or social workers with clinical specialties (15.7%, $n = 8$), some with extensive experience in treating refugee patients, and many with limited experience with this patient group. Length of treatment was flexible, at the most 1 to 2 years in the clinics, and without specific time limits in the individual practices.

Measures

The Rorschach components

We used the two pretreatment RIM components of the index study to investigate the aims and test the hypotheses of the study. The Trauma Response component, varying from constricted (−) to flooded (+), was understood as a dimension of personality functioning comprising degrees of complexity in thinking, emotional modulation problems, fear of forces outside personal control, traumatic intrusions, vigilance toward other people, and traumatic thought disorder (captured by Blends, CF+C, m, TCI, M, and SevCog, positively weighted, and F%, negatively weighted). Reality Testing, varying from impaired (−) to adequate (+), was understood as a perception based dimension comprising extent of faulty perception, ability to see features of the environment the way others see them, and notice prominent details and multiple aspects of the environment (captured by FQo, D, and R, positively weighted, and FQ−%, negatively weighted). The strength and relevance of the

components led us to expect that they might predict outcome of treatment and concurrent life events during follow-up.

The Harvard Trauma Questionnaire

The Harvard Trauma Questionnaire (HTQ; Mollica, McDonald, Massagli, & Silove, 2004) was developed to assess refugee trauma. We used Part I (HTQ−Trauma Events), which is a checklist of potentially traumatic experiences, and Part IV, Questions 1 to 16, symptoms of PTSD, derived from the *DSM−IV* criteria for PTSD. Symptoms for the last week are ranked on a 4-point scale ranging from 1 (*not at all*) to 4 (*extremely*). To compute mean scores we required 80% of the questions completed. We computed the mean score of the 16 items for the PTSD sum score (PTSD−Total), and the mean of scores on each *DSM−IV*-based symptom cluster; PTSD−Reexperiencing (Questions 1, 2, 3, and 16), PTSD−Arousal (Questions 6, 7, 8, 9, and 10), and PTSD−Avoidance (Questions 4, 5, 11, 12, 13, 14, and 15). An HTQ interview with a PTSD score of ≥2.5 is considered positive for PTSD. In research with traumatized refugees, internal consistency for HTQ Part IV, Questions 1 to 16, has been reported with Cronbach alphas ranging from .74 to .95 (e.g., Bentley, Thoburn, Stewart, & Boynton, 2012; Jakobsen, Thoresen, & Johansen, 2011; Kleijn, Hovens, & Rodenburg, 2001). We obtained a Cronbach's alpha of .85 at T1, .89 at T2, and .93 at T3. In most social science research, a Cronbach's alpha value of .70 is considered acceptable, and .80 is good (Fayers & Machin, 2007).

The Hopkins Symptom Checklist−25

The Hopkins Symptom Checklist−25 (HSCL−25; Mollica et al., 2004) assesses anxiety (10 items) and depression (15 items). The HSCL−25 has been found to have good reliability and validity in clinical refugee samples (e.g., Hollifield, Warner, & Lian, 2002). Symptoms for the last week are ranked on a 4-point scale from 1 (*not at all*) to 4 (*extremely*). To compute mean scores we required 80% of the questions completed. An anxiety score of > 1.75 represents a clinical level of anxiety that was found to be consistent with several anxiety diagnoses, whereas a depression score > 1.75 represents a clinical level of depression found to be consistent with major depressive disorder (Mollica et al., 2004). In studies of traumatized refugee populations of various ethnic backgrounds, internal consistency for the HSCL−25 has been reported with Cronbach's alpha ranging from .83 to .96 (e.g., Bentley et al., 2012; Renner, Salem, & Ottomeyer, 2006; Teodorescu et al., 2012). We obtained a Cronbach's alpha for the HSCL−25 of .91 at T1, .94 at T2, and .94 at T3.

The HTQ and the HSCL−25 have been widely used internationally in the assessment of mental disorders in traumatized populations (e.g., Jakobsen et al., 2011; Mollica et al., 2004) and as measures of treatment outcome (e.g., Carlsson et al., 2010; Hinton et al., 2004).

The World Health Organization Quality of Life−BREF

The World Health Organization Quality of Life−BREF (WHOQOL−BREF; WHOQOL Group, 1998) has 26 questions divided into four domains: physical health (Domain 1), psychological health (Domain 2), social relationships (Domain 3), and environmental conditions (Domain 4). The WHOQOL−BREF has frequently been used in research of traumatized populations, often in conjunction with the HTQ and the HSCL−25 (e.g., Teodorescu et al., 2012). Scoring is marked on 5-point Likert scales ranging from 1 (*very dissatisfied*) to 5 (*very satisfied*), *disagree strongly* to *agree strongly,* and related formulations. Computation of mean scores required completion of a minimum 6 of 7 items, 5 of 6 items, 2 of 3 items, and 6 of 8 items on Domains 1, 2, 3, and 4, respectively. We used the manual's conversion of raw scores to a 0 to 100 scale (*very poor* to *very good* [QOL]).

In a large international field trial by Skevington, Lotfy, and O'Connell (2004) among adults recruited from different sociodemographic backgrounds in the population and from physical and mental health clinical settings, the WHOQOL−BREF was found to be a valid QOL measure. The mean QOL values in Skevington et al.'s study were 64.8, 60.0, 57.2, and 54.0 for Domains 1, 2, 3, and 4, respectively. The internal consistencies for Domains 1, 2, and 4 (consisting of 6−8 items each) were reported with Cronbach's alpha values of .80 or more, and for Domain 3 (3 items only) it was .68 (Skevington et al., 2004). The Cronbach's alpha is sensitive to number of items (fewer than 10), and alpha values must then be interpreted with caution (Cortina, 1993). In this study we obtained $\alpha = .67$ for Domain 1, .76 for Domain 2, .59 for Domain 3, and .79 for Domain 4 at T1; at T2 we obtained .87, .77, .70, and .59; and at T3 we obtained .78, .76, .68, and .62 on the four domains, respectively.

Real-life functioning

The following variables were recorded at T1 and again at T3, and scored yes (1) or no (0) for the situation at each point in time: *Communicated in Norwegian,* determined by whether the participant was able to take part in the research interviews without an interpreter, and *Presently Employed,* determined by whether the participant held a job, part time or full time, even if on temporary sick leave. These indications of real-life functioning were crude. Being employed and learning the Norwegian language, or not, could be related to many factors. However, as unemployment in Norway is low, and language training is obligatory, we regarded getting work or starting to communicate in Norwegian as signs of healthier functioning, possibly cognitively, emotionally, and socially.

Variables related to treatment

We calculated *Frequency of Sessions* (number of sessions per month of treatment) and *Treatment Length* (in months) from information in the treatment records and from therapists' reports at T2 and T3. In addition, we recorded terminations, new periods of treatment, and any new treatment plans.

Procedures

In their written informed consent to take part in the study, the participants also allowed us to interview their therapists and access information about their treatment. The study was approved by the Norwegian Regional Committee for Medical and Health Research Ethics (REK, South-East; see https://helseforskning.etikkom.no) and adhered to the Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects (World Medical Association, 2010).

Treatment followed the Norwegian health laws and the ethical principles of the professions involved.

All patients were reassessed at T2 and T3, whether still in treatment or not, with a semistructured, qualitative interview and the HSCL−25, Part IV of the HTQ, and the WHOQOL−BREF. Interpreters were used when needed. The first and last authors administered all interviews and questionnaires, allowing time for participants to give associations, examples, and to ask for clarifications in response to the questions. Treatment charts and interviews with therapists provided us with details about treatment. Due to the participants' limited endurance, the numbers that completed the different questionnaires at follow-up varied.

Statistical procedures

To investigate the relationship between the RIM components, we correlated absolute component scores of Trauma Response with ordinary Reality Testing component scores. In addition, a scatterplot of the two components was drawn, including a smooth lowess (locally weighted scatterplot smoothing) curve, which showed that Trauma Response values around zero corresponded with positive Reality Testing scores, whereas flooded Trauma Response scores approximately above 1.0 and constricted Trauma Response scores below −0.5 corresponded with increasingly impaired Reality Testing. We used the mean absolute value of the "turning points" of the Trauma Response component on the scatterplot (−0.5 and 1.0) to divide our participants into three subgroups. Specifically, participants with Trauma Response component scores < -0.75 were labeled constricted, participants with Trauma Response scores between −0.75 and 0.75 were labeled middle (centered around our sample mean), and participants with Trauma Response scores >0.75 were labeled flooded. We used analysis of variance to compute subgroup means of the RIM variables constituting the component, and compared the means with international norms (Meyer, Erdberg, & Shaffer, 2007). We used the same procedure with the Reality Testing component scores, and used the same component values (−0.75 and 0.75) to divide participants into impaired, middle (around sample mean), and adequate Reality Testing subgroups. The division into subgroups was performed to evaluate the meaning of different component scores compared to RIM variable norms; otherwise, the continuous component scores were used in the analyses.

We used linear mixed effects modeling in a hierarchical setup to study relationships of the RIM components with the continuous outcome variables (symptoms and QOL). Interpretation of the estimated fixed effects coefficients in linear mixed effects models is analogous to interpretation of coefficients in linear regression models, in which regression coefficients are interpreted as adjusted slopes for continuous independent variables and adjusted differences between categories for categorical independent variables. We calculated change in percentage points (PP) per unit change in the predictor variable (100 × estimate/range of the scale), to allow comparison of the amount of change between differently scaled outcome variables (symptoms scaled 1−4, QOL scaled 0−100).

In our mixed effects models, the RIM components were modeled as continuous and time (T1, T2, and T3) was modeled as categorical. As the sample size did not allow investigation of nonlinear relationships, we used a technical procedure to permit different linear relationships of the Trauma Response component below and above zero. This was achieved by splitting Trauma Response into a positive and negative part. In the positive part, negative values were set to be zero, and in the negative part, all positive values were set to be zero. In the first stage of the mixed effects analysis, separate models included the relationship between each RIM component and the outcome variables, adjusted for time. Next, we included both RIM components and their interaction in a model with no time variable. Finally, in the mixed effects analyses, separate models included interactions between each component and time, comprising contrasts between T1 and T2, T2 and T3, and T1 and T3. Due to the sample size, it was not possible to include all effects in a unified model. We chose not to regard p values $< .05$ of the total model as an absolute requirement for considering a statistic effect as potentially real (Altman, 1991, pp. 168−169). We were concerned with characteristics of the process over time, as well as the total model. Along with one significant total model, we therefore included two models with significant time contrasts only, for further elaboration and discussion. Significant interactions and possible trends were visualized by a figure. We fitted a model where component scores were set to −1, 0, and 1, respectively. The figure was fashioned to illustrate that the time span from T1 to T2 was 1 year, and from T2 to T3 twice as long (2 years).

We used logistic regression to analyze the relationships between each RIM component and the dichotomous real-life variables Communicated in Norwegian and Presently Employed. As the degrees of freedom did not allow splitting of the Trauma Response component, we here used the absolute value of Trauma Response to account for the potential effects of deviations from zero in both directions. Finally, by linear regression in a hierarchical setup, we analyzed the relationships between the RIM components and the continuous variables Frequency of Sessions and Treatment Length, using the split Trauma Response component and Reality Testing separately and in interaction.

We used a method developed by Nakagawa and Schielzeth (2013) for calculating marginal R^2 as a measure of effect size in general linear mixed effects models. Marginal R^2 is defined as the proportion of variance attributed to the fixed factors, compared with the sum of variances for random effects. Further, we computed an effect size measure r defined as the square root of the marginal R^2. The strength of the relationships in the logistic regressions was estimated by odds ratios (OR), and is also reported as the square root of Nagelkerke's R^2. To estimate effect sizes in the linear regressions, we used partial eta (computed as the square root of partial eta squared), regarded as commensurate with r. According to Cohen's (1992) benchmarks, a Pearson correlation coefficient $r = .10$ represents a small effect, $r = .30$ a medium effect, and $r = .50$ a large effect. The linear mixed effects analyses were estimated using R (The R Foundation for Statistical Computing, Vienna, Austria, 2013; nlme package), and remaining analyses were run in SPSS (IBM SPSS Statistics, Version 20, 2011).

Results

Descriptives

Table 1 shows that from T1 to T3, the number of participants who could communicate in Norwegian increased, the number who were employed increased, all QOL measures increased, and all symptom measures decreased. However, the participants still suffered from high levels of mental health symptoms at follow-up, and comorbidity was still high. At T3, 65.9% still concurrently qualified for PTSD, major depressive disorder, and a clinical level of anxiety, compared to 81.6% at T1. Moreover, the mean QOL measures were still low at T3 compared with international means (Skevington et al., 2004). An exception was QOL-Environmental Conditions, where the score was closer to the international mean, probably indicative of the relatively good welfare system in Norway. Only one participant lived in a refugee reception center; the others lived in ordinary flats in the community. See Table 1 for details.

The diagnoses given in the clinics also demonstrated the comorbidity of these patients. According to treatment charts and written information from the therapists, 82.4% ($n = 42$) of the patients were diagnosed with the ICD-10 diagnoses PTSD (F 43.1) or enduring personality change after catastrophic experience (F 62.0) and 5.9% ($n = 3$) with other reactions to severe stress (F43) as main diagnoses, accompanied by various comorbid anxiety-related, depressive, dissociative, and somatoform disorders. Among the 51 patients, 5.6% had been given four diagnoses, 14.8% three, 57.4% two, and 22.2% had been given only one diagnosis.

There was great variation in frequency, length, and continuity of treatment. Within the first year, mean frequency of sessions per month of treatment was 1.8 ($SD = 1.4$) and mean treatment length was 10.6 months ($SD = 2.5$). At T2, 37.3% ($n = 19$) of the treatments were terminated. Only three patients had dropped out without notice or against the therapists' advice. For the entire follow-up period, mean Frequency of Sessions was 1.7 ($SD = 1.1$) and mean treatment length was 24.5 months ($SD = 13.7$). At T3, 29% ($n = 15$) were still in treatment and 14% ($n = 7$) who had terminated at some time during follow-up had an appointment to start in treatment again.

Characteristics of the RIM components

When using absolute values of the Trauma Response component, the correlation between the components was significant and of medium effect size ($r = -.30$, $p = .035$), showing that degree of Reality Testing difficulties and degree of flooding or constriction were not independent of each other. The scatterplot described in the Statistical Procedures section indicated a relationship between poorer Reality Testing and more "extreme" values of Trauma Response. These findings confirmed our hypothesis that extreme values of Trauma Response in both directions were related to poorer Reality Testing.

As can be seen in Table 2, means of the individual RIM variables constituting the Reality Testing component approached international norms with increasing component scores and were similar to international norm means in the high, adequate subgroup. The means of the Trauma Response component also confirmed our assumptions reasonably well: The mean values of most of the RIM variables constituting Trauma Response in the middle subgroup were close to the international norm means, and far below or above in the constricted and flooded subgroups. Exceptions to these findings were with the TCI and the SevCog, which were closer to international norms in the low, constricted subgroup and increasingly elevated with higher component scores.

Table 1. Real-life functioning, symptoms, and quality of life, at Time 1 (T1), Time 2 (T2), and Time 3 (T3).

	T1 (treatment start)			T2 (1-year follow-up)			T3 (3-year follow-up)		
Variable	*N*	*%/M*	*n/SD*	*N*	*%/M*	*n/SD*	*N*	*%/M*	*n/SD*
Indications of real-life functioning									
Communicated in Norwegian	51	50.9%	26[a]				50	70.0%	35
Presently Employed	51	21.6%	11				50	34.0%	14[b]
Indications of psychological disorder									
Anxiety > 1.75, indicating clinically significant anxiety	50	96.0%	48	46	89.1%	41	45	86.7%	39
Depression > 1.75, T1, indicating major depression	50	98.0%	49	46	91.3%	42	45	91.1%	41
PTSD-Total ≥ 2.5, indicating PTSD	49	81.6%	40	20	75.0%	15	41	65.9%	27
Symptom scores									
Anxiety	50	2.87	0.59	46	2.68	0.67	45	2.65	0.63
Depression	50	2.94	0.52	46	2.81	0.61	45	2.74	0.64
PTSD-Total	49	2.82	0.47	20	2.88	0.53	41	2.64	0.69
PTSD-Reexperiencing	49	2.91	0.68	20	2.84	0.67	49	2.75	0.78
PTSD-Arousal	49	3.10	0.49	20	3.13	0.47	41	2.86	0.78
PTSD-Avoidance	49	2.57	0.58	20	2.72	0.67	41	2.43	0.68
Quality of life (QOL)									
QOL-Physical Health	49	28.2	13.8	17	36.1	20.0	42	38.3	17.7
QOL-Psychological Health	49	25.3	15.7	17	30.0	15.1	42	33.5	16.6
QOL-Social Relationships	49	35.0	21.7	17	44.4	25.8	42	42.4	21.4
QOL-Environmental Conditions	49	45.0	18.5	17	44.6	14.9	42	48.5	14.1

Note. Anxiety = Questions 1–10, Depression = Questions 11–25, Hopkins Symptom Checklist–25 (HSCL–25), scaled 1–4. PTSD-Total = Questions 1–16; PTSD-Reexperiencing = Questions 1, 2, 3, and 16; PTSD-Arousal = Questions 6, 7, 8, 9, and 10; PTSD-Avoidance = Questions 4, 5, 11, 12, 13, 14, and 15, Harvard Trauma Questionnaire (HTQ), Part IV, scaled 1–4. QOL-Physical Health, QOL-Psychological Health, QOL-Social Relationships, and QOL-Environmental Conditions = Domains 1, 2, 3, and 4 of the World Health Organization Quality of Life–BREF (WHOQOL–BREF), scaled 1–100. For comparison, the international means on WHOQOL–BREF (Skevington et al., 2004) were 64.8, 60.0, 57.2, and 54.0 on Domains 1 to 4, respectively. T1 measures of HTQ, HSCL–25, and WHOQOL–BREF were presented in Opaas and Hartmann (2013).

[a]Two of the remaining participants communicated in English and did not need an interpreter. [b]Four individuals started to work and one gave up working during follow-up.

Table 2. RIM variable means in subgroups with low, medium, and high component scores.

	Trauma Response subgroups[a]							
	Constricted ($n = 12$)		Middle ($n = 29$)		Flooded ($n = 10$)		International norms	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	M_{int}	SD_{int}
Trauma Response score	**−1.20**	.30	**−.01**	.45	**1.47**	.68		
Blends	**0.33**	.49	**3.62**	2.06	**7.30**	2.87	**4.01**	2.97
CF+C	**0.50**	.67	**2.45**	1.48	**4.90**	1.60	**1.99**	1.88
m	**0.33**	.89	**1.86**	1.90	**6.00**	2.91	**1.50**	1.54
F%[b]	**.68**	.20	**.41**	.18	**.16**	.12	**.39**	.17
TCI	**.14**	.17	**.42**	.25	**.84**	.40	**.17**	.13
M	**0.42**	.52	**2.90**	1.76	**6.40**	3.81	**3.73**	2.66
SevCog	**.25**	.45	**1.72**	1.51	**4.40**	2.50	**0.5**	0.9
	Reality Testing subgroups[a]							
	Impaired ($n = 11$)		Middle ($n = 28$)		Adequate ($n = 12$)		International norms	
	M	*SD*	*M*	*SD*	*M*	*SD*	M_{int}	SD_{int}
Reality Testing score	**−1.53**	.51	**.10**	.42	**1.16**	.23		
FQ-%[b]	**.37**	.11	**.17**	.07	**.13**	.07	**.10**	.07
FQo%[b]	**.33**	.12	**.46**	.10	**.62**	.08	**.60**	.15
D	**1.82**	2.23	**6.32**	3.40	**11.0**	3.79	**9.89**	5.81
R	**13.45**	5.41	**17.57**	5.99	**22.50**	7.73	**22.31**	7.90

Note. RIM = Rorschach Inkblot Method. Means are shown in bold. M_{int}= means of the international norm data; SD_{int} = standard deviations of the norm data. Trauma Response and Reality Testing: RIM principal components (Opaas & Hartmann, 2013). Blends, CF, C, m, M, F%, D, R (Comprehensive System; Exner, 2003); TCI (Trauma Content Index; Armstrong & Loewenstein, 1990); and R−PAS (Meyer et al., 2011) variables SevCog (Severe Cognitive Codes: unweighted sum of Level 2 Cognitive Codes + CONTAM and ALOG), FQo% (Ordinary Form Quality%), and FQ-% (Form Quality Minus%). International norms from Meyer, Erdberg, and Shaffer (2007), Meyer et al. (2011; SevCog, FQo%, FQ-%), and Meyer (for Opaas & Hartmann, 2013; CF+C, TCI).
[a]The subgroups: RIM component scores < −0.75 (Low), −0.75 to 0.75 (Middle), and > 0.75 (High). [b] Stated as proportions.

Relationships of RIM components with symptoms, QOL, and real-life functioning

The mixed effects analysis, adjusted for time (Table 3), revealed a significant and negative relationship between Reality Testing and Anxiety, and between Reality Testing and PTSD-Arousal, and a significant, positive relationship between Reality Testing and QOL-Psychological Health. The Reality Testing component, adjusted for time, explained 9% to 13% of the variance in these outcome variables (r = .29, .31, and .35), resulting in a reduction of 5.3 and 5.0 percentage points on the symptoms scale and a 4.2 point increase in QOL-Psychological Health per unit change in Reality Testing. There were no other significant relationships. This implies that more adequate Reality Testing significantly predicted less Anxiety, less PTSD-Arousal symptoms, and higher QOL-Psychological Health throughout the study. Between the two components, adjusted for time, there were no significant interaction effects ($p \geq .27$) on symptoms and QOL.

Table 4 shows the overall interaction effects of RIM components and time on symptoms and QOL. For PTSD-Avoidance the interaction between the Reality Testing component and time was significant ($p = .016$, $r = .34$). Most other *p* values for the models were high, with no specific indications of overall interaction effects. In addition to the model involving PTSD-Avoidance, two other models had significant interaction with the time contrast T2 versus T1.

Table 5 shows details of the three interaction models with significant values overall (PTSD-Avoidance) or with significant interaction between Reality Testing and the time contrast T1 versus T2 (PTSD-Avoidance, PTSD-Reexperiencing, and PTSD-Total). Figure 1 illustrates the model-based differences in outcome variables over time for these three models when component scores were set to −1, 0, and 1, respectively. The models indicate that adequate Reality Testing (component score = 1) predicted a decrease in symptoms of PTSD the first year (in Avoidance, Reexperiencing, and Total symptoms), and retention of this improvement from T2 to T3. In contrast, impaired Reality Testing (component score = −1) predicted increasing symptoms of PTSD the first year (in PTSD-Avoidance symptoms, Reexperiencing, and Total symptoms) and decreasing symptoms from T2 to T3, but only back approximately to the T1 level.

According to the logistic regressions, a higher score on Reality Testing significantly and positively increased the chance that the participant communicated in Norwegian at T1 (OR = 2.00 [1.06, 3.77], $p = .033$, $r = .36$) and at T3 (OR = 2.28 [1.16, 4.51], $p = .017$, $r = .41$). A higher score on Reality Testing also significantly and positively increased the chance that the participant was employed at T1 (OR = 3.71 [1.27, 10.88], $p = .017$, $r = .48$) but not significantly at T3 (OR = 1.77 [0.85, 3.66], $p = .126$, $r = .27$). The absolute value of Trauma Response was nonsignificantly related with the real-life outcome variables ($p \geq .34$). Thus, increasing values of Reality Testing increased the chance of communicating in Norwegian at T1 and T3 and of being employed at T1.

Participants with more impaired Reality Testing, compared to those with adequate Reality Testing, had more mental health symptoms, had poorer QOL, and were less likely to communicate with us in Norwegian throughout follow-up, supporting our hypothesis. This was demonstrated by Anxiety and PTSD-Arousal symptoms, QOL in the psychological health domain, and in the proportion that communicated in Norwegian, but not Depression, other PTSD clusters, other QOL domains, or

Table 3. Relationships of RIM components with symptoms and QOL, adjusted for time.

	Estimate [95% CI]	PP	*t*(*df*)	*p*	R^2	*r*
Anxiety						
Trauma Response, negative values	−0.15 [−0.49, 0.19]	−5.0	−0.88 [1, 48]	.38	.06	.25
Trauma Response, positive values	0.22 [−0.05, 0.49]	7.3	1.64 [1, 48]	.11		
Reality Testing	−0.16* [−0.31, −0.01]	−5.3	−2.17* [1, 49]	.035	.09	.29
Depression						
Trauma Response, negative values	−0.04 [−0.37, 0.29]	−1.3	−0.23 [1, 48]	.82	.03	.18
Trauma Response, positive values	0.11 [−0.15, 0.37]	3.7	0.89 [1, 48]	.38		
Reality Testing	−0.06 [−0.21, 0.08]	−2.0	−0.88 [2, 49]	.38	.03	.18
PTSD-Total						
Trauma Response, negative values	0.09 [−0.22, 0.41]	3.0	0.58 [1, 48]	.56	.04	.21
Trauma Response, positive values	0.07 [−0.17, 0.31]	2.3	0.58 [1, 48]	.57		
Reality Testing	−0.13† (-0.27, 0.01)	−4.3	−1.86†[1, 49]	.069	.08	.28
PTSD-Reexperiencing						
Trauma Response, negative values	0.24 [−0.16, 0.63]	8.0	1.22 [1, 48]	.23	.05	.23
Trauma Response, positive values	0.07 [−0.23, 0.38]	2.3	0.48 [1, 48]	.63		
Reality Testing	−0.10 [−0.28, 0.09]	−3.3	−1.06 [1, 49]	.29	.03	.17
PTSD-Arousal						
Trauma Response, negative values	0.01 [−0.33, 0.35]	0.3	0.07 [1, 48]	.95	.04	.20
Trauma Response, positive values	0.07 [−0.19, 0.34]	2.3	0.57 [1, 48]	.57		
Reality Testing	−0.15* [−0.30, −0.01]	−5.0	−2.08* [2, 49]	.043	.09	.31
PTSD-Avoidance						
Trauma Response, negative values	0.11 [−0.23, 0.45]	3.7	0.45 [1, 48]	.66	.03	.18
Trauma Response, positive values	0.06 [−0.21, 0.32]	2.0	0.44 [1, 48]	.66		
Reality Testing	−0.11 [−0.27, 0.04]	−3.7	−1.68 [1, 49]	.099	.06	.25
QOL-Physical Health						
Trauma Response, negative values	6.85 [−2.33, 16.04]	6.9	1.50 [1, 48]	.14	.12	.34
Trauma Response, positive values	−5.40 [−12.70, 1.90]	−5.4	−1.49 [1, 48]	.14		
Reality Testing	3.92† [−0.16, 7.99]	3.9	1.93† [1, 49]	.059	.13	.36
QOL-Psychological Health						
Trauma Response, negative values	−0.85 [−9.85, 8.15]	−0.9	−0.19 [1, 48]	.85	.07	.26
Trauma Response, positive values	−1.74 [−8.90, 5.42]	−1.7	−0.49 [1, 48]	.63		
Reality Testing	4.21* [0.37, 8.06]	4.2	2.20* [1, 49]	.032	.13	.35
QOL-Social Relationships						
Trauma Response, negative values	−1.60 [−13.50, 10.30]	−1.6	−0.27 [1, 48]	.79	.06	.25
Trauma Response, positive values	−5.46 [−14.93, 4.00]	−5.5	−1.16 [1, 48]	.25		
Reality Testing	4.69† [−0.61, 9.99]	4.7	1.78† [1, 49]	.081	.08	.28
QOL-Environmental Conditions						
Trauma Response, negative values	6.51 [−2.36, 15.37]	6.5	1.48 [1, 48]	.15	.04	.21
Trauma Response, positive values	−1.92 [−8.97, 5.12]	−1.9	−0.55 [1, 48]	.59		
Reality Testing	3.49†[−0.46, 7.43]	3.5	1.78† [1, 49]	.082	.05	.23

Note. RIM = Rorschach Inkblot Method; QOL = quality of life. Linear mixed effects analysis. Trauma Response and Reality Testing = RIM principal components (Opaas & Hartmann, 2013). Estimate (fixed effects coefficient) = difference in outcome for a one-unit increase in a RIM component. PP = Estimate transformed to percentage points = 100 × estimate/range of the outcome scale. R^2 was computed by Nakagawa and Schielzeth's (2013) method for obtaining marginal R^2 from generalized linear mixed effects models. *r* = square root of the marginal R^2. Negative and positive values of Trauma Response were included in the same model.
*$p < .05$. †$p < .10$.

the proportion that was presently employed. The corresponding hypothesis for participants with highly flooded or highly constricted Trauma Responses was not supported. The results indicate a relationship between Reality Testing and the direction, amount, and path of change in symptoms of PTSD from T1 to T2 to T3, but did not demonstrate associations of any of the components with changes in other variables.

Relationships between RIM components and therapy "dosage"

According to the linear regression, Reality Testing was significantly and positively related, with medium effect sizes, with Frequency of Sessions during the first year ($t = 2.67$, $p = .010$, partial $\eta = .36$) and during all 3 years of follow-up ($t = 3.43$, $p = .001$, partial $\eta = .45$). The flooded half of Trauma Response (positive part) was subsignificantly and positively related to Frequency of Sessions the first year ($t = 1.81$, $p = .076$, partial $\eta = .25$). The interaction between Reality Testing and the flooded half of Trauma Response was subsignificant and positive for the 3 years of follow-up ($t = 1.79$, $p = .081$, partial $\eta = .26$). There were no other significant or subsignificant relationships of Trauma Response and Reality Testing, separately or in interaction, with Frequency of Sessions ($p \geq .25$), or with Treatment Length ($p \geq .31$).

Discussion

Two strong RIM components, Trauma Response and Reality Testing, characterized the trauma-related personality functioning of our sample of traumatized refugees at treatment start. In our further investigation of these components, we found that higher, positive scores of the Reality Testing component corresponded with values of the constituent RIM variables that were close to the international norms, indicating normal, healthy personality functioning in the associated areas of personality. Accordingly, lower component scores corresponded with less normal or healthy personality functioning. Derived from the interpretations of the individual RIM variables, adequate Reality Testing (positive component scores) indicated ability to

Table 4. Overall interaction effects of RIM components and time on symptoms and QOL.

	Overall interaction			
	F value (*df*)	*p*	R^2	*r*
Anxiety				
Time × Trauma Response, negative values	0.21 (2, 84)	.81	.06	.25
Time × Trauma Response, positive values	0.14 (2, 84)	.87		
Time × Reality Testing	0.75 (2, 86)	.48	.09	.30
Depression				
Time × Trauma Response, negative values	0.06 (2, 84)	.94	.04	.20
Time × Trauma Response, positive values	0.88 (2, 84)	.42		
Time × Reality Testing	0.07 (2, 86)	.93	.03	.18
PTSD-Total				
Time × Trauma Response, negative values	0.85 (2, 53)	.43	.06	.24
Time × Trauma Response, positive values	0.12 (2, 53)	.89		
Time × Reality Testing	*2.67*† *(2, 55)*	*.078*	*.11*	*.33*
PTSD-Reexperiencing				
Time × Trauma Response, negative values	0.38 (2, 53)	.69	.06	.24
Time × Trauma Response, positive values	0.20 (2, 53)	.82		
Time × Reality Testing	*2.29 (2, 55)*	*.11*	*.05*	*.23*
PTSD-Arousal				
Time × Trauma Response, negative values	0.55 (2, 53)	.58	.06	.25
Time × Trauma Response, positive values	0.93 (2, 53)	.40		
Time × Reality Testing	0.90 (2, 55)	.41	.10	.32
PTSD-Avoidance				
Time × Trauma Response, negative values	1.28 (2, 53)	.29	.05	.23
Time × Trauma Response, positive values	0.06 (2, 53)	.94		
Time × Reality Testing	*4.48** *(2, 55)*	*.016*	*.12*	*.34*
QOL-Physical Health				
Time × Trauma Response, negative values	0.14 (2, 51)	.87	.12	.35
Time × Trauma Response, positive values	0.10 (2, 51)	.91		
Time × Reality Testing	0.50 (2, 53)	.61	.13	.36
QOL-Psychological Health				
Time × Trauma Response, negative values	0.44 (2, 51)	.65	.07	.27
Time × Trauma Response, positive values	0.05 (2, 51)	.95		
Time × Reality Testing	0.29 (2, 53)	.75	.12	.35
QOL-Social Relationships				
Time × Trauma Response, negative values	2.55† (2, 51)	.088	.10	.31
Time × Trauma Response, positive values	0.29 (2, 51)	.75		
Time × Reality Testing	1.58 (2, 53)	.22	.09	.30
QOL-Environmental Conditions				
Time × Trauma Response, negative values	0.87 (2, 51)	.42	.05	.23
Time × Trauma Response, positive values	0.70 (2, 51)	.50		
Time × Reality Testing	0.91 (2, 53)	.41	.06	.25

Note. RIM = Rorschach Inkblot Method; QOL = quality of life. Mixed effects analyses of RIM components interacting with time. R^2 was computed by Nakagawa and Schielzeth's (2013) method for obtaining marginal R^2 from generalized linear mixed effects models. *r* = square root of the marginal R^2. Negative and positive values of the Trauma Response component are included in the same model for each outcome variable, resulting in one R^2value for each. Negative and positive values of Trauma Response were included in the same model. Models with significant overall values or significant time contrasts, detailed in Table 5 and visualized in Figure 1, are shown in italics.
$^{*}p < .05$. $^{\dagger}p < .10$.

accurately perceive and recognize obvious, composite, and multiple aspects of the environment, whereas lower scores indicated faulty perceptions and limited ability to recognize obvious, contextual, and multiple aspects of the perceptual field.

Further, middle values of the Trauma Response component corresponded with values on the constituent RIM variables close to international norms, thus associated with normal, healthy functioning. Higher positive or lower negative scores, indicating more severe flooding or constriction, corresponded with less normal or less healthy personality functioning in the associated areas. Exceptions among the RIM variables constituting this component were the TCI and SevCog, which were closer to the international norms in the low, highly constricted group and increasingly elevated with higher scores. This means that, except for those few who were severely cognitively and emotionally constricted, almost all our participants showed RIM evidence of traumatic intrusions (TCI) and disturbance of formal thinking (SevCog) above normal values, consistent with findings from other traumatized samples (e.g., Kaser-Boyd & Evans, 2008).

When using ordinary component scores in the index study, Trauma Response and Reality Testing had been relatively independent. However, the medium-size correlation between the components when using absolute scores of Trauma Response might indicate a curvilinear relationship. Viglione et al. (2012) posed several questions for further research regarding the RIM responses of traumatized individuals; for example, whether perceptual and thought disturbances were associated with overall complexity and thematic richness. We found that individuals with a highly flooded Trauma Response (indicating complexity, thought disturbances, and emotional control problems)—but also individuals with a highly constricted (impoverished) Trauma Response—showed less adequate Reality Testing (indicating perceptual disturbances).

We expected that the participants' scores on the components might be relevant to their progress through treatment and follow-up. We found that participants with adequate—compared to more impaired—Reality Testing throughout the study had better mental health (significantly less Anxiety and PTSD-Arousal symptoms), better QOL (higher score on QOL-Psychological Health), better real-life functioning (significantly more who communicated in Norwegian), and a more regular attendance at therapy (significantly higher Frequency of Sessions). These results could not be attributed to treatment, because the lower level of distress and better QOL were apparent at treatment start. We believe these results were sustained by more stable functioning in general in these patients. The relationship between Reality Testing and PTSD-Arousal was also negative at treatment start, but did not reach significance until follow-up. The significant relationships of Reality Testing with the QOL domains of Physical Health and Social Relationships, found in the index study, were not significant at follow-up.

Contrary to our expectations, we found no significant relationships between Trauma Response and any of the outcome variables. In the index study, a more flooded Trauma Response was significantly associated with more PTSD-Reexperiencing symptoms, but at follow-up, the relationship was nonsignificant. This indicates that participants' changes in PTSD-Reexperiencing during follow-up happened irrespective of their "position" on the Trauma Response component. It could appear as if Trauma Response had little association with the participants' mental health and functioning, and with their capacity to improve.

One explanation of the weak predictive capacity of the Trauma Response component could be the varying clinical presentations of traumatized individuals, fluctuating between flooded and constricted responses when exposed to traumatic reminders (e.g., Ephraim, 2002; Kaser-Boyd & Evans, 2008). Many of the individual variables constituting the Trauma Response component of this study are among the more state-like, less temporally stable, RIM variables in the study of Grønnerød (2003). Grønnerød's (2003, 2004) studies support

Table 5. Changes in symptoms and QOL related to interaction effects of RIM components and time, details.

	Difference between T1 and T2			Difference between T2 and T3			Difference between T1 and T3		
	Estimate [95% CI], *t*	PP	*p*	Estimate [95% CI], *t*	PP	*p*	Estimate [95% CI], *t*	PP	*p*
PTSD-Total									
Time × Reality Testing	−0.26* [−0.50, −0.02], −2.15	−8.7	.041	0.12 [−0.13, 0.37], 0.95	4.0	.34	−0.14 [−0.32, 0.05], −1.50	−4.7	.14
PTSD-Reexperiencing									
Time × Reality Testing	−0.31* [−0.60, −0.01], −2.10	−10.3	.041	0.28 [−0.03, 0.58], 1.80	9.3	.077	−0.03 [−0.26, 0.20], −0.27	−1.0	.79
PTSD-Avoidance									
Time × Reality Testing	−0.39**[−0.67, −0.11], −2.83	−13.0	.006	0.20 [−0.09, 0.49], 1.36	6.7	.18	−0.20†[−0.41, 0.02], −1.83	−6.7	.073

Note. QOL = quality of life; RIM = Rorschach Inkblot Method. Mixed effects analyses of RIM components interacting with time. Interaction models with significant values overall (PTSD-Avoidance) or with significant time contrasts (T1−T2). See Figure 1 for visualization. Degrees of freedom = 2, 55. Estimate = interaction coefficient in the mixed effects model. PP = Estimate transformed to percentage points = 100 × estimate/range of the outcome scale.
*$p < .05$. **$p < .01$. †$p < .10$.

the interpretation of Trauma Response as a more fluid personality dimension, more amenable to change and less likely to predict future relationships. Variables constituting the Reality Testing component seemed to be more trait-like, according to Grønnerød's (2003) findings.

Analyses of changes in symptoms and QOL associated with the RIM components showed significant interaction effects of Reality Testing in the period T1 to T2 in PTSD-Total, PTSD-Avoidance, and PTSD-Reexperiencing. There were no significant interaction effects with Trauma Response from T1 to T2, and no interaction effects of any component from T1 to T3. This means that we could not demonstrate that differences in Reality Testing or Trauma Response were associated with extent of improvement during the 3 years. One of the reasons for this was probably that improvement did not happen smoothly and incrementally from T1 to T2 to T3. As indicated by the models in Figure 1, different and partly opposite trajectories of change seemed to characterize subgroups among the participants. According to the model, patients with adequate Reality Testing improved markedly the first year, and, although the paths of change differed somewhat among the PTSD symptom clusters, these participants retained their overall improvement in symptoms of PTSD from T2 to T3. Patients with only slightly deficient Reality Testing (around sample mean) improved gradually, but modestly. Patients with impaired Reality Testing showed a marked increase in symptoms of PTSD from T1 to T2 and a decrease from T2 to T3, approximately back to the T1 level. The trajectories in anxiety, depression, and QOL were not significantly affected by the participants' "position" along the Reality Testing or Trauma Response personality dimensions.

Frequency of Sessions was positively and significantly associated with Reality Testing throughout follow-up, which indicated that patients with more adequate Reality Testing met more frequently for treatment. However, length of treatment was not significantly related with the patients' degree of Reality Testing or flooded or constricted Trauma Response. We found no significant relationships between treatment length and the components, and we did not find significant relationships between frequency or length of treatment and improvement.

The investigated changes in symptoms and QOL represented the interaction effects of the personality dimensions and time. In this case, time not only included the treatment received during follow-up, but also other life events happening during this time. We found that adequate Reality Testing was important to the continued and significantly lower psychiatric

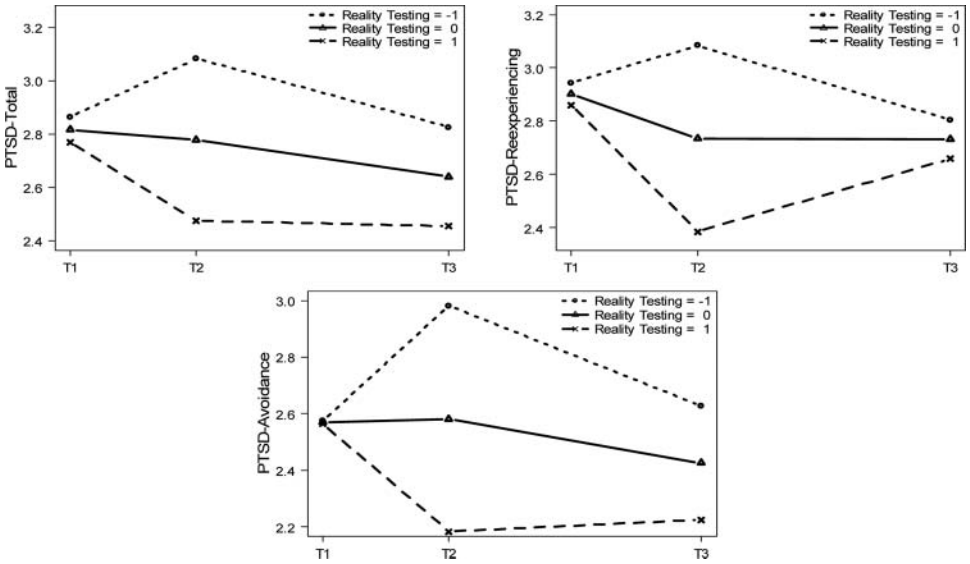

Figure 1. A model illustrating interaction effects over time of the Reality Testing component on outcome variables PTSD-Total, PTSD-Reexperiencing, and PTSD-Avoidance when component values are set to −1 (impaired Reality Testing), 0 (sample mean/somewhat lowered Reality Testing), and +1 (adequate Reality Testing).

symptomatology, better QOL, and Norwegian communication skills. The modeled paths of change throughout the 3 years suggest that improvement had different trajectories depending on whether Reality Testing scores were high, medium, or low. Some of the progress in participants with poorer Reality Testing did not show up until the second or third year of treatment, suggesting that treatment of these patients takes time. The total extent of change or improvement from treatment start to the 3-year follow-up did not significantly relate with any of the RIM components.

Most of the participants had RIM scores indicating disturbances of perception, Reality Testing, and logical thinking. Still, in our extended contact with them, only a few expressed psychotic-like ideas during interviews. Furthermore, none of the patients were diagnosed with a psychotic disorder by their therapists. Available research suggests that in posttraumatic conditions, traumatic intrusions disrupt thinking, judgment, concentration, and Reality Testing, but these functions remain intact when associated with nontrauma thoughts (e.g., Viglione et al., 2012). Although our data do not allow causal interpretations, these findings together with theories and findings within developmental psychology and neuropsychology (e.g., Fonagy & Target, 2005; Schore, 2002) provisionally guide our suggestion that the observed impaired Reality Testing and the elevated level of thought disturbances (SevCog) among many of our participants should be understood as primarily trauma-based.

Such impairment might be harder to detect than more predominant breaches in formal and logical reasoning. The therapeutic relationship is an arena where therapists can discover context-dependent misperceptions and faulty logic, although it might take time to identify. RIM assessment is another, faster method, that might give valid assessment before therapy starts. Rorschach-based assessment provides, to our knowledge, the only means of obtaining normatively based information concerning the conventionality of one's perceptions.

There could be several reasons why personality functioning, in terms of the two RIM components, did not have a greater impact on outcome. Patient characteristics interact with other variables in treatment in a dynamic process (Clarkin & Levy, 2004), making it unlikely to find simple and significant relationships of therapy variables or patient variables, with outcome. In a small sample like ours, this would be even harder. Furthermore, influences from treatment interact with the patients' living conditions, relationships, and external events.

Limitations

Our results must be interpreted with caution due to several limitations. The low sample size did not allow controlling for premorbid symptom level and examining interactions between both RIM components and time in a unified model. Furthermore, the low sample size prevented us from conducting curvilinear analyses of the Trauma Response variable, which would have been more appropriate than the technical procedure of splitting the sample. Our participants almost uniformly had high values on anxiety and depression, and a high and limited range on posttraumatic symptoms. With little variation and low sample size, we might have missed potentially important relationships. The study involved relationships that have not previously been addressed in research, which might generate new hypotheses and be included in meta-analyses. We therefore adopted a liberal stance to p values in reporting mixed effects models, including potential trends with significant time contrasts only. This must be taken into consideration when interpreting the results.

Other limitations were the variable response to the questionnaires and the possible misunderstandings and loss of information due to language barriers. Through the first and last authors' personal administration and presence through all data collection, we could observe that the research procedures to a fair degree seemed to mitigate these barriers. We responded to nonverbal expressions indicating difficulties in comprehension, allowed time for mutual clarifications, and welcomed the participants' associations and examples.

A reasonable concern is whether the assessments put too much strain on a vulnerable group. Several participants referred to our long, repeated, and clinically informed research interviews as part of their treatment. Our impression was that most of our participants experienced the assessment as therapeutic. This probably contributed positively to the participants' experience of taking part in research, and to the low dropout rate. At the same time, the research interviews represented an important deviation from ordinary care.

One advantage of the naturalistic design was that it increased the external validity of the results and the generalizability to ordinary refugee patients. In this respect, the wide inclusion criteria, the high response and inclusion rate, and the low dropout rate of the study, even among patients who dropped out of or terminated treatment, were strengths of our study.

Implications

Our model showed a rapid improvement in symptoms of PTSD in participants with adequate Reality Testing, across various therapists, treatment sites, and theoretical orientations. This could indicate that patients with adequate Reality Testing can benefit from various therapeutic approaches. On the other hand, the modeled increase in symptoms of PTSD the first year among patients with impaired Reality Testing, and their slight improvement, suggest a need to study under which conditions such patients might improve. Based on this study, for patients with impaired Reality Testing, we recommend efforts at securing regular attendance to therapy and a focus on how they perceive the therapy and the realities around them. Further, helping highly constricted patients to develop language and other symbolic expressions of their experiences and emotions, and highly flooded patients to tolerate trauma reminders, will probably enhance regulation of emotions, and perhaps improve Reality Testing. Follow-up periods in research should be long enough to detect changes. The RIM components found to be characteristic of this sample of traumatized refugees might extend our understanding of the psychology of traumatization and be of use in the continuing work toward finding the best therapeutic approach for the individual patient along the therapeutic process. Comparable studies are needed to validate and extend these results.

Acknowledgments

Marianne Opaas presented preliminary parts of this study at the annual meeting of the Society of Personality Assessment, Arlington, VA, in 2014. We would like to thank patients, therapists, and mental health clinics for their participation.

Funding

This research was fully financed by the Norwegian Centre for Violence and Traumatic Stress Studies, Oslo, Norway.

References

Allen, J., & Dana, R. H. (2004). Methodological issues in cross-cultural and multicultural Rorschach research. *Journal of Personality Assessment*, *82*, 189–208.

Altman, D. G. (1991). *Practical statistics for medical research*. London, UK: Chapman & Hall/CRC.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Armstrong, J. G., & Loewenstein, R. J. (1990). Characteristics of patients with multiple personality dissociative disorders on psychological testing. *Journal of Nervous and Mental Disease*, *178*, 448–454.

Arnon, Z., Maoz, G., Gazit, T., & Klein, E. (2011). Rorschach indicators of PTSD: A retrospective study. *Rorschachiana*, *32*, 5–26.

Bentley, J. A., Thoburn, J. W., Stewart, D. G., & Boynton, L. D. (2012). Post-migration stress as a moderator between traumatic exposure and self-reported mental health symptoms in a sample of Somali refugees. *Journal of Loss and Trauma: International Perspectives on Stress & Coping*, *17*, 452–469.

Boehnlein, J. K., Kinzie, J. D., Sekiya, U., Riley, C., Pou, K., & Rosborough, B. (2004). A ten-year treatment outcome study of traumatized Cambodian refugees. *Journal of Nervous and Mental Disease*, *192*, 658–663.

Bohart, A. C., & Wade, A. G. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 219–257). Hoboken, NJ: Wiley.

Bornstein, R. F. (2012). Rorschach score validation as a model for 21st-century personality assessment. *Journal of Personality Assessment*, *94*, 26–38.

Brand, B. L., Armstrong, J. G., & Loewenstein, R. J. (2006). Psychological assessment of patients with dissociative identity disorder. *Psychiatric Clinics of North America*, *29*, 145–168.

Carlsson, J. M., Olsen, D. R., Kastrup, M., & Mortensen, E. L. (2010). Late mental health changes in tortured refugees in multidisciplinary treatment. *Journal of Nervous and Mental Disease*, *198*, 824–828.

Clarkin, J. F., & Levy, K. N. (2004). The influence of client variables on psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp 194–226). NewYork, NY: Wiley.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98–104.

Courtois, C. A. (2008). Complex trauma, complex reactions: Assessment and treatment. *Psychotherapy: Theory, Research, Practice, Training*, *41*, 412–425.

Crumlish, N., & O'Rourke, K. (2010). A systematic review of treatments for post-traumatic stress disorder among refugees and asylum-seekers. *Journal of Nervous and Mental Disease*, *198*, 237–251.

Ephraim, D. (2002). Rorschach trauma assessment of survivors of torture and state violence. *Rorschachiana*, *25*, 58–76.

Evans, B. F. (2008). The Rorschach and immigration evaluations. In C. B. Gacono & F. B. Evans (Eds.), *The handbook of forensic Rorschach assessment* (pp. 489–504). New York, NY: Taylor & Francis.

Exner, J. E., Jr. (2003). *The Rorschach: A comprehensive system: Vol.1* (4th ed.). New York, NY: Wiley.

Fayers, P. M., & Machin, D. (2007). *Quality of life: The assessment, analysis and interpretation of patient-reported outcomes*. Chichester, UK: Wiley.

Fonagy, P., & Target, M. (2005). Bridging the transmission gap: An end to an important mystery of attachment research? *Attachment and Human Development*, *7*, 333–343.

Fowler, J. C., Ackerman, S. J., 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 Rorschach, TAT, and DSM–IV Axis V. *Journal of Personality Assessment*, *83*, 306–322.

Graceffo, R. A., Mihura, J. L., & Meyer, G. J. (2014). A meta-analysis of an implicit measure of personality functioning: The Mutuality of Autonomy Scale. *Journal of Personality Assessment*, *96*, 581–595.

Grønnerød, C. (2003). Temporal stability in the Rorschach method: A meta-analytic review. *Journal of Personality Assessment*, *80*, 272–293.

Grønnerød, C. (2004). Rorschach assessment of changes following psychotherapy: A meta-analytic review. *Journal of Personality Assessment*, *83*, 256–276.

Hartmann, E., & Hartmann, T. (2014). The impact of exposure to Internet-based information about the Rorschach and the MMPI–2 on psychiatric outpatients' ability to simulate mentally ealthy test performance, *Journal of Personality Assessment*, *96*, 432–444.

Herman, J. L. (1992). Complex PTSD: A syndrome in survivors of prolonged and repeated trauma. *Journal of Traumatic Stress*, *5*, 377–391.

Hilsenroth, M. J., Handler, L., Toman, K. M., & Padawer, J. R. (1995). Rorschach and MMPI-2 indices of early psychotherapy termination. *Journal of Consulting and Clinical Psychology*, *63*, 956–965.

Hinton, D. E. & Kirmayer, L. J. (2013). Local responses to trauma: Symptom, affect, and healing. *Transcultural Psychiatry*, *50*, 607–621.

Hinton, D. E., Pham, T., Tran, M., Safren, S. A., Otto, M. W., & Pollack, M. H. (2004). CBT for Vietnamese refugees with treatment-resistant PTSD and panic attacks: A pilot study. *Journal of Traumatic Stress*, *17*, 429–433.

Hollifield, M., Warner, T. D., & Lian, N. (2002). Measuring trauma and health status in refugees: A critical review. *Journal of the American Medical Association*, *288*, 611–621.

Holt, R. (1977). A method for assessing primary process manifestations and their controls in Rorschach responses. In M. Rickers-Ovsiankina (Ed.), *Rorschach psychology* (pp. 375–420). New York, NY: Krieger.

Jakobsen, M., Thoresen, S., & Johansen, L. E. E. (2011). The validity of screening for post-traumatic stress disorder and other mental health problems among asylum seekers from different countries. *Journal of Refugee Studies*, *24*, 171–186.

Kaser-Boyd, N., & Evans, F. B. (2008). Rorschach assessment of psychological trauma. In C. B. Gacono & F. B. Evans (Eds.), *The handbook of forensic Rorschach assessment*(pp. 255–277). New York, NY: Taylor & Francis.

Kleijn, W. C., Hovens, J. E., & Rodenburg, J. J. (2001). Posttraumatic stress symptoms in refugees: Assessments with the Harvard Trauma Questionnaire and the Hopkins Symptom Checklist–25 in different languages. *Psychological Reports, 88*, 527–532.

Klopfer, B., Kirkner, F. J., Wisham, W., & Baker, G. (1951). Rorschach prognostic rating scale. *Journal of Projective Techniques, 15*, 425–428.

Lambert, M. J. (2013). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 169–218). Hoboken, NJ: Wiley.

Levin, P., & Reis, B. (1997). Use of the Rorschach in assessing trauma. In J. P. Wilson &. T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 529–543). New York, NY: Guilford.

McFarlane, C. A., & Kaplan, I. (2012). Evidence-based psychological interventions for adult survivors of torture and trauma: A 30-year review. *Transcultural Psychiatry*, *49*, 539–567.

Meyer, G. J., Erdberg, P., & Shaffer, T. W. (2007). Towards international normative reference data for the Comprehensive System. *Journal of Personality Assessment*, *89*(Suppl. 1), S201–S216.

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. (2008). An introduction to Rorschach assessment. In R. P. Archer & S. R. Smith (Eds.), *Personality assessment* (pp. 281–336). New York, NY: Routledge.

Meyer, G. J., Viglione, D. J., & Exner, J. E., Jr. (2001). Superiority of Form% over Lambda for research on the Rorschach Comprehensive System. *Journal of Personality Assessment*, *76*, 68–75.

Meyer, G. J., Viglione, D. J., Mihura, J., Erard, R. E., & Erdberg, P. (2011). *Rorschach Performance Assessment System: Administration, coding, interpretation, and technical manual.* Toledo, OH: Rorschach Performance Assessment System.

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). A systematic review and meta-analysis of the Rorschach Comprehensive System validity literature. *Psychological Bulletin*, *139*, 548–605.

Mollica, R. F., McDonald, L. S., Massagli, M. P., & Silove, D. M. (2004). *Harvard program in refugee trauma: Measuring trauma, measuring torture. Instructions and guidance on the utilization of Hopkins Symptom Checklist–25 (HSCL–25) & the Harvard Trauma Questionnaire* (HTQ).Retrieved from http://hprt-cambridge.org/

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142.

Norcross, J. C., & Lambert, M. J. (2011). Evidence-based therapy relationships. In J. C. Norcross (Ed.), *Psychotherapy relationships that work. Evidence-based responsiveness*(2nd ed., pp. 3–21). New York, NY: Oxford University Press.

Opaas, M., & Hartmann, E. (2013). Rorschach assessment of traumatized refugees: An exploratory factor analysis. *Journal of Personality Assessment*, *95*, 457–470.

Perry, W., & Viglione, D. J. (1991). The Ego Impairment Index as a predictor of outcome in melancholic depressed patients treated with tricyclic antidepressants. *Journal of Personality Assessment*, *56*, 487–501.

Renner, W., Salem, I., & Ottomeyer, K. (2006). Cross-cultural validation of measures of traumatic symptoms in groups of asylum seekers from Chechnya, Afghanistan, and West-Africa. *Social Behaviour and Personality*, *34*, 1101–1114.

Schore, A. N. (2002). Dysregulation of the right brain: A fundamental mechanism of traumatic attachment and the psychopathogenesis of posttraumatic stress disorder. *Australian and New Zealand Journal of Psychiatry*, *36*, 9–30.

Schottenbauer, M. A., Glass, C. R., Arnkoff, D. B., & Gray, S. H. (2008). Contributions of psychodynamic approaches to treatment of PTSD and trauma: A review of the empirical treatment and psychopathology literature. *Psychiatry*, *71*, 13–34.

Shaffer, T. W., Erdberg, P., & Meyer, G. J. (2007). Introduction to the *JPA* special supplement on international reference samples for the Rorschach Comprehensive System. *Journal of Personality Assessment*, *89* (Suppl. 1), S2–S6.

Skevington, S. M., Lotfy, M., & O'Connell, K. A. (2004). The World Health Organization's WHOQOL–BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL Group. *Quality of Life Research*, *13*, 299–310.

Smith, J. D., Van Ryzin, M. J., Fowler J. C., & Handler L. (2014). Predicting response to intensive multimodal inpatient treatment: A comparison of single- and multiple-class growth modeling approaches. *Journal of Personality Assessment*, *96*, 306–315.

Steel, Z., Silove, D., Phan, T., & Bauman, A. (2002). Long-term effect of psychological trauma on the mental health of Vietnamese refugees resettled in Australia: A population-based study. *The Lancet*, *360*, 1056–1062.

Stokes, J. M., Pogge, D. L., Powell-Lunder, J., Ward, A. W., Bilginer, L., & DeLuca, V. A. (2003). The Rorschach Ego Impairment Index: Prediction of treatment outcome in a child psychiatric population. *Journal of Personality Assessment*, *81*, 11–19.

Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *80*, 547–559.

Teodorescu, D., Siqveland, J., Heir, T., Hauff, E., Wentzel-Larsen, T., & Lien, L. (2012). Posttraumatic growth, depressive symptoms, posttraumatic stress symptoms, post-migration stressors and quality of life in multi-traumatized psychiatric outpatients with a refugee background in Norway. *Health and Quality of Life Outcomes*, *10*, 1–16.

Urist, J. (1977). The Rorschach test and the assessment of object relations. *Journal of Personality Assessment*, *41*, 3–9.

Vaage, A. B., Thomsen, P. H., Silove, D., Wentzel-Larsen, T., Van Ta, T., & Hauff, E. (2010). Long-term mental health of Vietnamese refugees in the aftermath of trauma. *British Journal of Psychiatry*, *196*, 122–125.

van der Kolk, B. A. (1994). The body keeps the score: Memory and the evolving psychobiology of posttraumatic stress. *Harvard Review of Psychiatry*, *1*, 253–265.

van der Kolk, B. A., & Ducey, C. P. (1989). The psychological processing of traumatic experience: Rorschach patterns in PTSD. *Journal of Traumatic Stress*, *2*, 259–274.

Van Ee, E., Kleber, R. J., & Mooren, T. T. M. (2012). War trauma lingers on: Associations between maternal posttraumatic stress disorder, parent–child interaction, and child development. *Infant Mental Health Journal*, *33*, 459–468.

van Wyk, S., & Schweitzer, R. D. (2014). A systematic review of naturalistic interventions in refugee populations. *Journal of Immigrant Minority Health*, *16*, 968–977.

Viglione, D. J. (1999). A review of recent research addressing the utility of the Rorschach. *Psychological Assessment*, *11*, 251–265.

Viglione, D. J., Perry, W., Giromini, L., & Meyer, G. J. (2011). Revising the Rorschach Ego Impairment Index to accommodate recent recommendations about improving Rorschach validity. *International Journal of Testing*, *11*, 349–364.

Viglione, D. J., Towns, B., & Lindshield, D. (2012). Understanding and using the Rorschach inkblot test to assess post-traumatic conditions. *Psychological Injury and Law*, *5*, 135–144.

Weiner, I. B. (2003). Principles of Rorschach interpretation. In I. B. Weiner (Series Ed.), *Personality and clinical psychology series* (2nd ed.). Mahwah, NJ: Erlbaum.

WHOQOL Group. (1998). Development of the World Health Organization WHOQOL–BREF quality of life assessment. *Psychological Medicine*, *28*, 551–558.

World Health Organization Geneva. (1994). *Pocket guide to the ICD-10 classification of mental and behavioural disorders, with glossary and diagnostic criteria for research, ICD-10: DCR-10*. London, UK: Churchill Livingstone, Elsevier.

World Medical Association. (2010). *WMA Declaration of Helsinki—Ethical principles for medical research involving human subjects.* Retrieved from http://www.wma.net/en/30publications/10policies/b3/17c.pdf