

2009

# An examination of the validity of the Rorschach Ego Impairment Index (EII-2) using the John Hopkins Precursors Study cohort

George Bombel  
*The University of Toledo*

Follow this and additional works at: <http://utdr.utoledo.edu/theses-dissertations>

---

## Recommended Citation

Bombel, George, "An examination of the validity of the Rorschach Ego Impairment Index (EII-2) using the John Hopkins Precursors Study cohort" (2009). *Theses and Dissertations*. 1041.  
<http://utdr.utoledo.edu/theses-dissertations/1041>

This Dissertation is brought to you for free and open access by The University of Toledo Digital Repository. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of The University of Toledo Digital Repository. For more information, please see the repository's [About page](#).

A Dissertation Entitled  
An Examination of the Validity of the Rorschach Ego Impairment Index (EII-2)  
Using the Johns Hopkins Precursors Study Cohort

By  
George Bombel

Submitted as partial fulfillment of the requirements for  
The Doctor of Philosophy Degree in Clinical Psychology

---

Advisor: Joni L. Mihura, Ph.D.

---

Gregory J. Meyer, Ph.D.

---

Jeanne H. Brockmyer, Ph.D.

---

John D. Jasper, Ph.D.

---

Steven Huprich, Ph.D.

---

Graduate School

The University of Toledo

August 2009

An Abstract of  
An Examination of the Validity of the Rorschach Ego Impairment Index (EII-2)  
Using the Johns Hopkins Precursors Study Cohort

George Bombel

Submitted as partial fulfillment of the requirements for  
The Doctor of Philosophy degree in Psychology

The University of Toledo

August 2009

The Rorschach Ego Impairment Index (EII / EII-2; Perry & Viglione, 1991; Viglione, Perry, & Meyer, 2003) was created to assess one's capacity to organize and utilize internal resources for coping with demands of both internal and external stressors. The scale is scored from Comprehensive System (CS; Exner, 2003) variables, and has good psychometric qualities (Perry & Braff, 1994; Perry, McDougall, & Viglione, 1995; Stokes et al., 2003); however, the use of behavioral criterion variables has been relatively infrequent in the EII validity literature (Cadenhead, Perry, & Braff, 1996; Perry & Braff, 1994). Further, few studies have investigated the incremental validity of the EII beyond self-report measures of distress, and none have examined its ability to predict long-term health outcomes in a prospective cohort (Dawes, 1999; Perry, 2001; Perry & Viglione, 1991; Stokes et al., 2003). The proposed study examined the predictive construct validity of the EII-2 using data from the prospective Johns Hopkins Precursors Study dataset. It

was hypothesized that EII-2 scores would predict the cumulative incidences of depression, mortality, and divorce over time; the cumulative incidence of depression and mortality over time would be significantly greater for high EII-2 scorers than low EII-2 scorers; EII-2 scores would predict later life psychological health; the interpersonal components of the EII-2 (*GHR*, *PHR*) would predict later life psychological health better than the thought disorder component (*WSum6*), which would predict better than a combined interpersonal / thought disorder component (*M-*), and all of these components would predict better than primitive contents; higher EII-2 scores (more ego impairment) would correlate significantly and negatively with later life physical health, the use of preventative health services, and perceived social support, but significantly and positively correlate with later life alcohol use, job dissatisfaction, and an aggregate index of all variables (except mortality) coded so that higher scores indicate more ego impairment. The only hypothesis that was supported was that the EII-2 correlated significantly and positively with the aggregated total impairment variable. Secondary analyses controlling for the possible effects of the original Rorschach examiners and protocol complexity did not alter findings for any variable. Possible reasons for the lack of significant findings in the survival and linear regression analyses were explored: (a) The EII-2 may be a better measure of ‘state’ manifestations of ego functioning as opposed to trait-level core personality structure; (b) the EII-2 may not be sensitive to the variability of ego functioning at the healthier end of the health-pathology spectrum; (c) sparse or absent inquiry material in many JHPS protocols may have depressed *WSum6* and EII-2 scores; (d) unexpected relationships among dependent variables suggested that some may have questionable validity; (e) time-to-first-event data used in survival analyses may not adequately capture how ego

functioning is expressed over time. Limitations of this study also include a design that limits generalizability of the findings to other populations.

## Acknowledgements

I am very grateful to Dr. Joni Mihura, my dissertation committee Chairperson, and Dr. Greg Meyer, committee member, for their encouragement, support, enthusiasm, and empathic attunement during this project, and throughout my graduate school experience. I am also grateful to committee members Drs. Jeanne H. Brockmyer, John D. Jasper, and Steven Huprich. They were patient and flexible, and provided helpful input. Knowingly or unknowingly, they also effectively modeled how to deal with stressed and frantic graduate students.

My sincerest thanks also go to Assessment Lab members Nick Katko, Aaron Upton, Nicolae Dumitrascu, Sandra Horn, and Elizabeth Koonce for their camaraderie, diligence, and intermittent flashes of bawdy humor. A special thanks goes to Nicolae for being so willing to score so many more protocols than were originally assigned to him.

Finally, I am grateful to Drs. Michael Klag and Dan Ford, and Lucy Meoni. This project would not have been possible but for their willingness to make the Johns Hopkins Precursors Study data available to the Assessment Lab. Their help was also invaluable during the delicate processes of hypothesis refinement and variable selection.

## Table of Contents

Abstract.....	ii - iv
Acknowledgements.....	v
Table of Contents.....	vi - vii
List of Figures.....	viii
List of Tables.....	ix
I. Introduction.....	1
Statement of the Problem.....	1
The Ego.....	4
Ego Assessment.....	7
Beres' Model.....	11
The Ego Impairment Index.....	15
Long-Term Prospective Research and the Precursors Study.....	28
Hypotheses.....	37
II. Method.....	41
Participants.....	41
Measures.....	43
Procedures.....	47
Analyses.....	49
Calculation of Self-Report and Aggregate Variables.....	51
Moderator Variables and Secondary Analyses.....	54
III. Results.....	56
Interrater and Scale Reliability.....	56

	Descriptive Statistics and Diagnostics.....	58
	Hypothesis 1: Predictive Validity (Depression).....	65
	Hypothesis 2: Predictive Validity (Depression).....	66
	Hypothesis 3: Predictive Validity (Psychological Health).....	67
	Hypothesis 4: Predictive Validity (Psychological Health).....	68
	Hypothesis 5: Predictive Validity (Mortality).....	70
	Hypothesis 6: Predictive Validity (Mortality).....	70
	Hypothesis 7: Predictive Validity (Divorce).....	71
	Hypotheses 8 – 13: Construct Validity.....	72
IV.	Discussion.....	78
V.	References.....	87



List of Figures

1. Ego Functions in Beres' Model and Corresponding EII/CS  
Variables.....19

2. Scatterplot of EII-2 Partial Residuals by Follow-Up Years in the Depression  
Analysis.....61

3. Scatterplot of EII-2 Partial Residuals by Follow-Up Years in the Mortality  
Analysis.....62

4. Scatterplot of EII-2 Partial Residuals by Follow-Up Years in the Divorce  
Analysis.....63

5. Scatterplot of Standardized Residual Values of EII-2 Components by the Standardized  
Predicted Scores for the SF-36 Psychological Health Scale.....65

6. Cumulative Hazard Function Plot of Depression by Follow-Up Time in Years for High  
and Low EII-2 Scorers.....67

7. Cumulative Hazard Function Plot of Depression by Follow-Up Time in Years for High  
and Low EII-2 Scorers.....71

List of Tables

1. EII Factor Score Coefficients.....20

2. EII/EII-2 Factor Score Coefficients.....26

3. Interrater Reliability Coefficients for EII-2 and Components.....57

4. Cronbach’s Alpha Coefficients for Self-Report Measures.....58

5. Descriptive Statistics for Study Variables.....59

6. Cox Regression Results for EII-2 as Predictor of Time to First Depressive Episode...66

7. Hierarchical Linear Regression Results for Total EII-2 Score as Predictor of SF-36  
Psychological Health Scale Score.....68

8. Hierarchical Linear Regression Results for EII-2 Components as Predictors of SF-36  
Psychological Health Scale Scores.....69

9. Cox Regression Results for EII-2 as Predictor of Time to Death.....70

10. Cox Regression Results for EII-2 as Predictor of Time to First Divorce.....72

11. Bivariate Correlations between the EII-2 and Self-Report and Aggregate Construct  
Validity Variables.....73

12. Intercorrelations Between Primary Study Variables.....75

## Chapter 1

### Introduction

#### *Statement of the Problem*

The Ego Impairment Index (EII; Perry & Viglione, 1991) was created to assess one's capacity to organize and utilize internal resources for coping with demands of both internal and external stressors. Although a non-Comprehensive System (CS; Exner, 2003) Rorschach scale, it was derived from five CS variables (*Sum FQ-*, *WSum6*, *M-*, *Human Experience Variable [HEV]*, and primitive contents) chosen by Perry and Viglione (1991) to reflect overlapping ego functions in Beres' (1956) model of ego assessment. These functions are reality-testing, thought disturbance, the quality of internalized object representations, defensive functioning, and the regulation and control of instinctual drives. Recently, the *HEV* was replaced in Exner's CS by the Human Representational Variable (*HRV*; Viglione, Perry, Jansak, Meyer, & Exner, 2003), which has better psychometric properties. Concomitantly, Viglione, Perry, and Meyer (2003) revised the EII by replacing the *HEV* with the *HRV*. However, they found that the factor score coefficients of the EII and the new EII-2 were highly correlated ( $r = .99$ ), and the properties of their distributions were almost identical. They concluded that the EII and EII-2 are essentially interchangeable (Viglione, Perry, & Meyer, 2003).

The EII appears to be quite stable over time (Perry, McDougall, & Viglione, 1995), and it has demonstrated excellent interrater reliability (Perry & Braff, 1994; Stokes

et al., 2003). Further, it has an impressive validity track record which suggests it can assess a trait-level, core personality construct that is relatively independent of presenting symptoms status, and subtle enough to be missed by other forms of assessment, e.g., self-report and observational (Perry & Viglione, 1991; Perry, Viglione, & Braff, 1992; Stokes et al., 2003). Relatively few studies in the EII literature have examined the predictive and incremental validities of the EII, though results have generally been impressive (Dawes, 1999; Perry, 2001; Perry & Viglione, 1991; Stokes et al., 2003). Therefore, a primary purpose of this study was to add substantially to this literature by utilizing longitudinal data from the Johns Hopkins Precursors Study, which was initiated by Dr. Caroline Bedell Thomas to investigate youthful precursors to later disease and death. Data collection began in 1946, and the cohort of subjects consists of the 17 Johns Hopkins medical school classes that graduated from 1948 through 1964 ( $N = 1,337$ ; Thomas, 1976). Baseline assessment was gathered from a variety of physical, physiological, psychosocial, and psychological measures. This included the Rorschach, which was administered to 937 cohort members in individual or group format by 10 different examiners using Beck's administration guidelines (Beck, 1944; Beck et al., 1961). Most of the surviving participants still complete annual follow-up questionnaires (Graves, Mead, Wang, Liang, & Klag, 1994; Torre et al., 2005).

In the current study, predictive validity was addressed by examining the extent to which EII-2 scores would predict the incidence of depression over time. Here, the Cox proportional hazards regression (Cox, 1972) was utilized. Further, a Kaplan-Meier analysis (Kaplan & Meier, 1958) was used to determine if the incidence of depression over time is greater among high EII-2 scores than low EII-2 scorers. A log rank test was

used to determine if the differences between the two groups were significant. Because psychiatric illness appears to increase the risk of mortality in general (Drew, 2005), and ego-impairment is theoretically related to psychopathology (Vaillant, 1994), another Cox proportional hazards regression was conducted to examine the extent to which EII-2 scores predict the incidence of mortality over time. A Kaplan-Meier analysis was run to determine if the incidence of mortality over time will be significantly greater for high EII-2 scores than low EII-2 scorers. Finally, because ego impairment is likely to negatively influence judgment, impulse control, perception of reality, and interpersonal relationship quality, another Cox regression was conducted to determine if EII-2 scores would predict the cumulative incidence of divorce in this cohort. The predictive validity of the EII-2 has never been examined using a long-term prospective cohort.

A hierarchical linear regression was conducted to determine if the overall EII-2 score would predict later life psychological health. Here, the EII-2 was the independent variable (IV) and the dependent variable (DV) was a self-report measure of psychological health administered in later life. The linear regressions were not repeated with mortality incidence as the DV because they require a DV that has continuous, interval properties, and each cohort member can only experience death one time. Importantly, because these analyses did not use time-to-event as the DV like the Cox regressions, they were significantly less powerful. Therefore, the EII-2 may significantly predict the cumulative incidence of depression using the Cox regression, but not be a significant predictor of psychological health in later life time using the linear regression.

Consistent with recommendations from previous authors (Perry & Viglione, 1991; Perry et al., 1992; Viglione, Perry, Jansak, et al., 2003), this study also examined the

relative contributions of the EII-2 subcomponents (*Sum FQ-*, *WSum6*, *M-*, *HRV*, and Primitive Contents) in predicting meaningful criterion variables. Here, the incremental validity of the EII-2 subcomponents to predict later life psychological health was examined with a hierarchical linear regression model. The Primitive Contents component was entered as block 3, *M-* was entered as block 4, *WSum6* was entered as block 5, and the interpersonal variables (*GHR*, *PHR*) were entered as the final block.

Finally, many studies in the EII literature have demonstrated convergent validity by relating EII scores to diagnostic groups, and/or to self or other-report measures of symptom severity (Adrian & Kaser-Boyd, 1995; Auslander, Perry, & Jeste, 2002; Cadenhead, Perry, & Braff, 1996; Perry, Minassian, Cadenhead, Sprock, & Braff, 2003; Perry, Moore, & Braff, 1995; Perry et al., 1992). In their article introducing the EII-2, Viglione et al. (2003) advocated including as criterion variables more behavioral markers of the constructs tapped by the EII-2, including indicators of information processing, decision making, and adaptive functioning. One purpose of this study, therefore, was to examine the construct validity of the EII-2 by relating the EII-2 overall score to follow-up questionnaire data regarding alcohol use, job dissatisfaction, perceived social support, physical health concerns, the use of preventive health services, and an aggregate variable based on all of these construct validity variables.

### *The Ego*

In English scientific literature, ‘ego’ antedates Freud. John Dewey and William James, for example, were using the term to refer to intentionality, or will, as early as the 1870s (Allport, 1943; James, 1879). By 1914, psychologists were using ‘ego’ as synonymous with mind or ‘self’ in psychological journals (Calkins, 1915; Dunlap, 1914).

Subsequent definitions of 'ego' have varied dramatically, and included ego as knower and intender, object of self-knowledge, the seat of primitive selfishness and impulses to dominate, passive organizer of mental processes, active behavioral organizer, and the subjective organizer of internalized culture (Allport, 1943).

Freud alluded to the 'ego' in his earliest works, e.g., *Studien über Hysterie* (*Studies on Hysteria*; Breuer & Freud, 1895/1957), but a formal definition of his famous construct did not reach English readers until 1927 when *Das Ich und Das Es* (*The Ego and the Id*) was translated by Joan Riviere (Freud, 1923/1927). In this work, Freud defined the ego as a structure which developed from the id specifically to mediate between the id's raw and immediate internal needs and the demands of the environment. The ego is born of this conflict, and circumscribed by it (Allport, 1943). To be successful at its task, it must (a) perceive both instinctual impulses and environmental circumstances, (b) intend and control body movements, and (c) decide if and when impulses should be expressed, or if they should even be permitted into consciousness (Monte, 1987). Later theorists would expand on this conception. Anna Freud pointed out that the id (instincts) wasn't the only source of conflict that caused the ego to mobilize defenses. The superego, in the form of guilt, and the ego itself, when it effectively perceives real threat, can create anxiety which must be contained to preserve the self. Further, Anna Freud re-framed our understanding of defense mechanisms to include the notion that they are not necessarily pathological, but necessary and healthy at times, especially for children (McWilliams, 1994).

Some theorists have argued, though, that the Freudian conception of the ego as a simple conflict mediator doesn't capture the complexity of the construct. In object

relations theory, the drive for relating to others is more prominent than the drive to satisfy instincts, and the ego is conceived as the executive controller of functions such as perception, learning, memory, and motor behavior. It also is responsible for organizing representations of self and others into an integrated whole (Christopher, Bickhard, & Lambeth, 2001; Kernberg, 1986). Thus, its role is not limited to finding a compromise between the instinctual needs and the environment.

In a similar vein, the expression of instinctual needs is not of primary importance in self psychology; rather, the id, ego, and superego are considered features of a more fundamental personality directed by the self (Patton, Connor, & Scott, 1982). Here, the ego operates in non-conflict spheres, as in object relations theory, but the self directs personality processes, including the ego-functions.

In Loevinger's (1976; Manners & Durkin, 2001) theory of ego development, the ego is the "master trait" which organizes and unifies the personality. As such, it regulates all other developmental processes. It is comprised of 4 domains: Character development, cognitive style, interpersonal style, and conscious preoccupations. The ego is said to "develop" as these domains are re-organized in response to the interpersonal and physical environments. Concomitantly, the individual moves through sequential, hierarchical, stages toward greater awareness, cognitive complexity, flexibility, autonomy, and responsibility (Manners & Durkin, 2001). Loevinger's ego is akin to Kohut's self.

Heinz Hartmann suggested that "...adaptation to reality...is directed by the organized structure of ego-functions (such as intelligence, perception, etc.) which exist in their own right and have an independent effect upon the solution of conflicts" (Hartmann, 1939, p. 214). Here, the ego's primary task remains adaptation; however, its role is no



longer that of passive mediator between internal needs and environmental demands, as in Freud's model. It is also motivated by energies that are not the product of internal conflict, such as biological maturation, learning, and the desires to organize, plan, understand, and master (Allport, 1943; Beres, 1956).

### *Ego Assessment*

Given its importance, the ego quickly became a focus in the personality assessment literature. Barron (1953) distinguished psychotherapy responders from non-responders after 6-months of treatment using 68 items from the Minnesota Multiphasic Personality Inventory (MMPI). Based on the content of these items, he concluded retrospectively that the new scale "...appears to measure the various aspects of effective personal functioning which are usually subsumed under the term 'ego-strength'" (Barron, 1953, p. 327). Accordingly, it was called the Ego-Strength (Es) Scale; it has been used to examine the relationship between ego-strength and variables such as anxiety and adjustment (Gottesman, 1959), defensiveness and coping behavior (Hunter & Goodstein, 1967; King & Schiller, 1960), inpatient and outpatient psychiatric treatment outcome (Adams & Cooper, 1962; Levine & Cohen, 1962; Sinnett, 1962), psychosocial adaptation to cancer (Worden & Sobel, 1978), locus of control (Sadowski, Woodward, Davis, & Elsbury, 1983), time perspective and the capacity to delay gratification (Shybut, 1970), anger and hostility (Schill & Thomsen, 1987), psychiatric diagnosis (Quay, 1963), the length of recovery from medical illness (Greenfield, Roessler, & Crosley, 1959), creativity (Dudek & Hall, 1984), and oral dependency (Oneill & Bornstein, 1990). Crumpton, Cantor, and Batiste (1960) extracted 13 meaningful factors from the Es, and concluded that the scale is sensitive to the presence of ego weakness, not ego strength.

Importantly, however, efforts to validate the Es scale as a predictor of psychotherapy response have produced contradictory findings. For example, higher pretreatment scores have been associated with more and less change during psychotherapy, while other data has indicated that there is no relationships between Es scale scores and psychotherapy response (Graham, 1987).

The Bell Object Relations Reality Testing Inventory (BORRTI; Bell, Billington, & Becker, 1985; Bell, Billington, & Becker, 1986) is as a self-report measure with 90 true-false items comprising a total of 7 scales assessing two domains of functioning. Four scales specifically target object relations (Alienation, Insecure Attachment, Egocentricity, and Social Incompetence), and three target reality-testing (Reality Distortion, Uncertainty of Perception, and Hallucinations and Delusions; Alpher, 1991). The BORRTI reality-testing scales have been used to investigate the relationship between ego-functioning and phenomena such as the onset of psychosis (Greig, Bell, Kaplan, & Bryson, 2000), and executive functioning in schizophrenia (Bell & Zito, 2005). It has demonstrated discriminant and construct validity (Alpher, 1991; Holaday & Glidewell, 2000); but, the Uncertainty of Perceptions subscale correlates significantly with social desirability (Bell et al., 1995), and the object relations subscales are moderately correlated among themselves (Buelow, McClain, & McIntosh, 1996).

Researchers have also used structured clinical interviews to assess ego function. Working from Beres' (1956) model of ego assessment, Leopold Bellak identified the following 12 ego functions during a 5-year National Institute of Mental Health (NIMH) study (Bellak & Hurvich, 1969): Reality-testing, judgment, sense of reality, regulation of drives, object relationships, thought processes, adaptive regression, defensive functioning,

stimulus barrier, autonomous functioning, synthetic functioning, and mastery-competence. He subsequently constructed a measure of these ego functions using over 24 rating scales to measure the combined effect of diazepam and psychotherapy (Bellak, Chassan, Gediman, & Hurvich, 1973). Allen, Coyne, and David (1986) used 10 of the Bellak scales identified through factor analysis to assess the relationship between ego-functioning and intelligence. Later, Gabbard et al. (2000) demonstrated that the scales were sensitive to therapeutic gains from intensive inpatient treatment for patients with severe personality disorders. Conte, Buckley, Picard, and Karasu (1995) created a measure (Self Evaluation Questionnaire) based on 4 of Bellak's 12 ego functions, and successfully used it as a criterion variable in a construct validity study of their Psychological Mindedness Scale (PMS; Conte et al., 1990).

Implicit personality assessment methods have long been used to assess underlying personality characteristics, or aspects of psychological functioning that operate unconsciously (Frank, 1939; Meyer & Kurtz, 2006). Based on Loevinger's (1976) theory of ego development described earlier, the Washington University Sentence Completion Test (WUSCT; Loevinger & Wessler, 1970) consists of 36 sentence stems which the test-taker is instructed to complete. Theoretically, the completed sentences betray the individual's level of ego development based on Loevinger's 9-stage theory. Sutton and Swenson (1983) obtained significant correlations between WUSCT scores and both an unstructured interview and TAT stories scored for ego development. In a factor analytic study supporting the unitary nature of Loevinger's ego concept, Loevinger and Wessler (1970) found that first and second factors accounted for 20% and 5.6% of the variance respectively, while the third and fourth factors were uninterpretable.

A number of systems have been created to score Rorschach responses for ego-strength or functioning. One of the oldest and most frequently used is the Rorschach Prognostic Rating Scale (RPRS; Klopfer, Ainsworth, Klopfer, & Holt, 1954). The RPRS was developed expressly as a predictor of response to psychotherapy, and as an index of available and potential ego resources. The variables in Klopfer's scoring system that contribute to the RPRS are Human, Animal, and Inanimate Movement, Shading – including Texture and Vista – Shading use problems, Color, Color use problems, and Form level, or quality (Meyer & Handler, 1997). The scale has been utilized to assess treatment outcome in an inpatient setting (Filmer-Bennett, 1955), investigate the relationships among implicit and self-report ego-strength assessment methods (Adams & Cooper, 1962), and demonstrate symptomatic improvement in an untreated, outpatient psychiatric sample (Endicott & Endicott, 1964). Cartwright (1958) used an abbreviated version of the RPRS to predict response to client-centered psychotherapy, and Benveniste, Papouchis, Allen, and Hurvich (1998) used Cartwright's version to examine the relationship between ego-functioning and annihilation anxiety. The RPRS has demonstrated excellent predictive validity; however, it is complicated to score, and Klopfer's scoring system is rarely taught today (Meyer & Handler, 1997).

Another popular Rorschach measure of ego-functioning is the Last-Weiss Ego-Strength Scale (Last & Weiss, 1976), which is composed of good form level variables from scoring systems by Exner (1986) and Klopfer (Klopfer et al., 1954). The variables are human movement ( $M+$ ), animal movement ( $FM+$ ), controlled color ( $FC+$ ,  $CF+$ ), and white space ( $S+$ ). Scores are summed to obtain the *Sum E* index, which indicates general adaptive capacity. Harder, Greenwald, Ritzler, Strauss, and Kokes (1988) found that,

surprisingly, higher *Sum E* values were correlated with negative outcomes for psychiatric inpatients at the 2-year follow-up. *S+* appeared to account for most of the variance in this relationship. The same year, Greenwald, Harder, Ritzler, Strauss, and Kokes (1988) obtained the opposite results with psychiatric outpatients: *Sum E* was significantly correlated with positive outcome measures. This time, only *M+* and *FC+* contributed to the relationships. Finally, Greenwald (1990), with a sample of 62 urban university students, found that the Last-Weiss Ego-Strength Scale did not correlate with the MMPI Es (Barron, 1953) described above.

Becker (1956) developed the Rorschach Genetic-Level (GL) Scoring System for adaptive functioning based on Werner's (1948) theory of development, and Beck's approach to administration and scoring. Axelrod and Kessel (1972) constructed a scale based on Weiner's (1966) Ego Disturbance Model for use with Rorschach responses scored according to Klopfer's system. The results from both of these studies were promising, but unfortunately, neither was cited after 1980, according to the Social Sciences Citation Index.

More recently, Perry and Viglione (1991) created the EII as a measure of adaptive functioning based on the ego assessment model of David Beres.

#### *Beres' Model of Ego Assessment*

Beres (1956) elaborated on Hartmann's conception to describe a cluster of ego-functions that overlap, but also perform unique adaptive roles. His model of ego assessment has been inspiring researchers for 50 years. It was the basis for the ego-disturbance model of assessment outlined in Weiner's classic 1966 book, *Psychodiagnosis of Schizophrenia* (Perry, 1994), and as noted, it informed Bellak's formulation of ego

functions and their assessment (Bellak & Hurvich, 1969). It has been integrated with family systems theory in the context of clinical social work assessment (Kaneko, 1984), and used to develop an interview-based measure to assess adaptive processes in adolescents (Beardslee et al., 1986). Fenchel and Flapan (1986) created a framework for understanding the therapeutic effects of analytic group psychotherapy based on Beres' work; and recently, Juni and Stack (2005) related addiction severity to a Beres-inspired measure of ego functioning.

In his monograph *Ego Deviation and the Concept of Schizophrenia*, Beres (1956) outlined seven ego-functions as follows:

*Relation to reality.* This function corresponds to the aspect of Freud's conception of the ego, which is charged with perceiving reality accurately. Theoretically, it develops gradually as the child becomes more responsible for fulfilling her own needs via interactions with the environment. Successful adaptation requires the ability to distinguish between fantasy and reality, and to transcend self-centeredness and an overly concrete notion of the 'objective' world (Beres, 1956). Respectively, neurosis and psychosis involve ignoring and denying reality in order to manage anxiety, with the latter representing a more severe degree of defensive departure that includes delusions and hallucinations (Freud, 1969). Manifestations in neurosis can include displacing emotions from threatening objects to safer objects (displacement), attributing one's own unacceptable emotions or desires onto another (projection), and barring unacceptable thoughts and feeling from consciousness via 'forgetting' (repression; Magnavita, 1997).

*Regulation and control of instinctual drives.* When ego-functions are adequately developed, the capacity to postpone or completely inhibit need gratification compliments

an accurate perception of reality. The individual is aware of his impulses, and can navigate inanimate and interpersonal environments to discharge them safely, or he can maintain safety by not discharging them at all. The mature ego has especially neutralized the sexual and aggressive impulses in accordance with societal standards (Freud, 1962; Hartmann, 1953). When ego-functions are impaired or underdeveloped, the relation to reality may or may not be disrupted; but in either case, the consistent capacity to inhibit impulses can be deficient. ‘Outbursts’ of this nature are typically diffuse, unrelated to any current frustration, and devoid of any goal except discharge. Mild outbursts can include thoughts or images that are not even detected as unacceptable by the ego (Beres, 1956).

*Object relationships.* Theoretically, the development of object relationships interdepends with the development of many other psychological functions, including reality-testing, identification, self-representation, and inhibition (Beres, 1956). Hartmann (1953) noted the healthy object relationships are only possible when an individual has been able to experience early caregivers as constant, and able to satisfy basic biological and emotional needs. Regression to narcissistic functioning, a constitutional inability to form attachments, and the mother’s inability to consistently provide need-satisfaction, can result in poor object relationships in children and adults (Hartmann, 1953; Mahler, 1952). Dynamically, in any case, this can appear as autistic withdrawal, attachment to inanimate or fantasy objects or animals, diffuse interpersonal clinginess, or prominent self-centeredness (Beres, 1956).

*Thought processes.* In psychoanalytic theory, thinking becomes more logical, organized, and rational as the psyche graduates from being dominated by primitive, autistic fantasy (primary process) to secondary processes, which permit effective

interactions with the environment. The capacity to reality-test as an adult, therefore, is intimately associated with the integrity of thought processes. Both are disrupted when primitive impulses intrude into consciousness. Creativity, slips of the tongue, and dreams are mild manifestations, while cognitive disorganization and bizarre thought content indicate more severe disruptions (Beres, 1956).

*Defense functions.* In turn, defense functions are intimately associated with reality-testing and thought organization. When the latter two functions are disrupted due to the emergence of unwanted unconscious material, the psyche chooses among available defensive strategies to dissolve the associated anxiety. According to Beres (1956), individuals with poor ego-functions utilize more primitive defense mechanisms such as projection, while healthier egos utilize mature defenses like sublimation or repression (for a more recent discussion of the hierarchy of defenses, see Vaillant, 1994). Further, through repetition, the use of some defenses can become habitual in certain situations (Freud, 1937). As noted, the failure of defensive functions can result in cognitive disorganization, or the intrusion of thought content that is bizarre or merely unacceptable.

*Autonomous functions* refer to the ego-functions described earlier which arise independently of the fundamental conflict between internal reality and the environment, but which may, nevertheless, become involved in this conflict. Hartmann (1939) included the following in a list of these autonomous ego-functions: Perception, intention, comprehension of objects, thinking, language, recall, motor development, grasping, crawling, walking, and other maturational and learning processes. Certain learning disorders, for example, can occur because of an underlying organic deficit, or as an



hysterical resolution to a longstanding internal conflict (Beres, 1956); or conceivably, for both reasons.

*The synthetic function* is the drive to organize, simplify, integrate, and understand. Beres (1956) and other theorists (e.g., see Hartmann, 1957, and Nunberg, 1931) have described it as the most primitive, comprehensive, and instrumental of the ego-functions because it can organize the others, and is the last stand against complete dissolution of the personality. Indeed, it can remain intact when other functions have failed and the individual no longer has adequate reality-testing. This permits some degree of adaptive functioning when thought content is bizarre, or thought processes are otherwise disorganized. An example would be an individual with schizophrenia who continues to meet her own basic needs for health and safety despite being homeless and communicating only in neologisms or 'word salad.' Synthesis is also involved when delusions have been integrated into an elaborate system, or when an individual progresses from primary to secondary process thinking. Confusion and the inability to function adequately in any sphere are symptomatic of the breakdown of the synthetic function (Beres, 1956).

#### *The Ego Impairment Index*

Perry and Viglione (1991) were interested in creating a scale that could assess personality and depression. They noted that, since self-report measures are subject to distortion, and focus only on symptomatology, they are unlikely to generate a full assessment picture. A measure of *underlying capacity*, therefore, might prove to be the best approach. The authors argued that the ambiguous Rorschach task is ideally suited to reveal underlying ego functioning because it requires that individuals use their own

internal psychological resources and organization to render the blots meaningful. They condensed Beres' (1956) seven ego-functions down to five, and identified CS variables which "...have been reported to...independently...delineate psychopathology across a variety of dimensions" (Perry & Viglione, 1991, p. 489). For example, *FQ-* (distorted form quality) responses mostly, or wholly, disregard blot properties, and often appear to be arbitrarily imposed (Exner, 2003). They are typically difficult for the examiner to find. On Card III, for example, the blot area at location D3 is commonly identified as "a butterfly" or "a pair of lungs" because the contours of the blot in this location are very consistent with the shape of a butterfly or a pair of lungs. However, "a red crab" would constitute form quality distortion because, although many crabs are reddish, the shape of this part of the blot is not consistent with the shape of a crab. High *Sum FQ-* values, therefore, indicate challenges with perceptual accuracy and reality-testing (Exner, 2003; Perry & Viglione, 1991).

When ego-functioning has been compromised, defensive functions like repression begins to fail, and unacceptable thoughts, feelings, wishes, and fantasies begin to break into consciousness. The presence of primitive content in Rorschach responses alludes to this failure (Beres, 1956). For the EII, Perry and Viglione (1991) chose the following primitive content categories described by Exner (1986) and Viglione (1990): *Anatomy (An)*, *Blood (Bl)*, *Explosion (Ex)*, *Fire (Fi)*, *Food (Fd)*, *Sex (Sx)*, *X-ray (Xy)*, and *Aggressive Movement (AG)*. An example of a response with primitive content from several categories would be: (Card III, Location W) "Two baboons tearing apart a baby gazelle for lunch. The red paint is blood." This notion of primitive content corresponds to Beres' (1956) instinct regulation and defense functioning categories.

Object relationships are estimated by Good Human Experience (*GHE*) and Poor Human Experience (*PHE*) – collectively, the Human Experience Variable (*HEV*) – as well as *M*-. *GHE* responses consist of percepts of whole humans with good form quality (i.e., accurately perceived), and the absence of otherwise tangential, peculiar, implausible, illogical, and/or bizarre verbal content. These percepts may be human “parts,” or even fictionalized humans (e.g., a giant) when the percept has been seen frequently by the norming groups. *GHE* can also be scored when cooperative movement (COP) is present among humans or animals, so long as the verbal content is not particularly bizarre or illogical, and doesn’t include aggression or morbidity. On Card II, the response “Two bears clapping hands together” would qualify as *GHE* (Perry & Viglione, 1991).

In contrast, *PHE* is scored for human content when (a) it is inaccurately perceived (i.e., poor *FQ*) or contains bizarre or illogical components; (e) it is part-human or fictionalized, and includes aggression or morbidity, or is not a cooperative or popular response (for more detailed scoring guidelines, see Perry & Viglione, 1991). An example of the latter response would be (Card 10, Location W) “A broad corset of penny-wigs, all dancing in a jumping unison formation!”

To obtain the *HEV*, which summarizes the schema of internalized relationships, *GHE* and *PHE* are transformed to *Z*-scores, and the former is subtracted from the latter (Perry & Viglione, 1991; Viglione, Perry, Jansak, Meyer, & Exner, 2003).

In addition to *FQ*-, primitive content, and the *HEV*, the EII includes human movement responses with poor *FQ* (*M*-) and a measure of disturbed thought processes. *M*-responses are rare, and generally hint at impaired thought processes as well as primitive internalized object representations (Exner, 2003; Perry & Viglione, 1991). Exner (1993)

noted that such responses are associated with deficient social skills, and poor interpersonal relationships. Further, M- responses contribute to the CS Perceptual Thinking Index (PTI; Exner, 2003). The response (Card I, Location W) "A person diving into some water" contains human movement, but the percept does not fit the contours of the blot.

Finally, the sum of the 6 weighted CS Cognitive Special Scores (*WSum6*) corresponds to Beres' disturbed thought processes; it indicates the degree to which responses in the protocol are tangential, peculiar, implausible, illogical, and/or bizarre (Exner, 2003). The 6 Special Scores are as follows: Deviant Verbalizations (*DV*) include neologisms or redundancies (e.g., "The dead corpse of a person."); Deviant Responses (*DR*) are inappropriate or circumstantial phrases (e.g., "It's a bat, but I wanted to see a butterfly."); Inappropriate Combinations (*INCOM*) are percepts where implausible properties or actions are attributed to a single object (e.g., "A man with a tail."); Fabulized Combination (*FABCOM*) includes implausible relationships between two or more objects (e.g., "A butterfly swallowing the HMS Pinafore."); Contamination (*CONTAM*) responses represent the fusion of two objects in an impossible way (e.g., "The face of a bug-ox."); and Inappropriate Logic (*ALOG*) include responses where unconventional reasoning is used to justify an answer without prompting (e.g., "This must be the North Pole because its at the top of the card.").

Two levels of Special Scores are included in the CS: Level 1 scores are "...mild or modest [instances] of illogical, fluid, peculiar, or circumstantial thinking" Exner, 2003, p. 135), whereas Level 2 scores are moderate to severe instances. The latter are usually easy to identify because of their bizarre or unusual nature. All Cognitive Special Scores except *CONTAM* and *ALOG* can be scored with both Levels. The *WSum6* value is obtained by

summing the products of the frequency of each special score in the record times a weight based on the severity of the problematic thinking represented by each score.

Perry and Viglione did not identify CS variables to target autonomous and synthetic functions in Beres' model. Figure 1 presents Beres' ego-functions and the corresponding CS variables in the EII.

Figure 1

*Ego-Functions in Beres' Model and Corresponding  
EII / CS Variables*

Beres' Model	EII / CS Variables
Relation to reality	<i>Sum FQ-, M-</i>
Instinct regulation	<i>An, Bl, Ex, Fi, Fd, Sx, Xy, AG</i>
Defense functions	"
Object relationships	Good Human Experience Poor Human Experience <i>M-</i>
Thought processes	<i>WSum6</i>

*Note.* EII = Ego Impairment Index; CS = Exner's (2003) Comprehensive System.

To compute an overall EII score, Rorschach responses are scored with the CS, and the values for each EII / CS variable are simply multiplied by their corresponding factor score coefficients, then summed. Perry and Viglione (1991) calculated these coefficients from a principle components analysis of Rorschach CS data obtained from a sample of depressed subjects (Haller, 1982). The EII / CS variables and their factor score coefficient values are in Table 1. Coefficients for the number of responses (*R*) and the constant were

included by Perry and Viglione in the final EII equation, so they are included here, although *R* was not found to affect factor loadings.

Table 1

*EII Factor Score Coefficients*

EII / CS Variables	Factor Score Coefficient
<i>Sum FQ-</i>	.136
Primitive Contents	.068
<i>GHE</i>	-.160
<i>PHE</i>	.108
<i>M-</i>	.208
<i>WSum6</i>	.050
<i>R</i>	-.061
Constant	-.049

*Note.* EII = Ego Impairment Index; CS = Exner's (2003) Comprehensive System.

After constructing the EII, Perry and Viglione (1991) sought to validate it. Consistent with initial hypotheses, low EII scores were associated with a more favorable response to tricyclic antidepressants among adult melancholic depressed patients ( $n = 46$ ). The criterion variables were the Beck Depression Inventory (BDI; Beck, 1967), a measure of cognitive aspects of depression, and the Carroll Rating Scale (CRS; Carroll, Feinberg, Smouse, Rawson, & Greden, 1981), which assesses more somatic symptoms. After demographic and baseline self-report measure data were partialled out of the regression analysis, the EII accounted for a significant amount of variance in treatment response. Further, when EII components were examined with a stepwise regression, only the *GHE* and *PHE* were significant predictors of BDI and CRS scores.

In a follow-up study to validate the EII on a sample of patients with schizophrenia, Perry, Viglione, and Braff (1992) hypothesized that (a) this new adult sample ( $n = 26$ ) would score significantly higher on the EII than the depressed sample in the initial study, (b) EII scores would be positively correlated with MMPI scales that assess psychosis (6, 8, and 9), and a measure of bizarre thought content (Magical Ideation Scale [MIS]; Eckblad & Chapman, 1983) and (c) because the disorganized and undifferentiated forms of schizophrenia are believed to involve less ego stability and integration than the paranoid type, a combined group of the former should score higher (more pathological) than a group of the latter. All hypotheses were supported, and factor score coefficients from a principle components analysis on the data from this sample correlated highly ( $r = .98$ ) with factor score coefficients from the initial study.

Perry and Braff (1994) went on to examine the relationship between information-processing impairments, as measured by behavioral procedures (e.g., visual backward masking [VBM], and a startle response prepulse inhibition) and thought disorder (EII) in schizophrenia. As hypothesized, higher EII scores were associated with poor information-processing performance in an adult sample ( $n = 52$ ). Importantly, the *PHE*, and not the *GHE*, component of the EII demonstrated significant correlations with the performance measures. Interrater reliabilities for the EII components were reported as Kappa coefficients ranging from .88 to .97.

Perry, Moore, and Braff (1995) investigated gender differences in information-processing and social competence among a sample of adults with schizophrenia ( $n = 87$ ). Based on literature suggesting that males have a more severe and disabling form of schizophrenia, the authors hypothesized that their male participants would demonstrate

more thought disorder on the dependent measures (i.e., score higher) than females. The criterion measures were the Scale for the Assessment of Positive Symptoms (SAPS), the Brief Psychiatric Rating Scale (BPRS), and the EII. Interestingly, the hypothesis was only confirmed with the EII, i.e., self-report measures of psychiatric symptoms did not distinguish males and females. Further, a rater-report measure of social functioning, the Social Competency Index (SCI), was also significantly correlated with thought disorder.

The same year, Perry, McDougall, and Viglione (1995) completed an extended study of the temporal stability of the EII. A very small sample ( $n = 17$ ) of the melancholic adults from the original EII study were re-assessed after 5 years, and rank-order correlations were computed between the two sets of EII scores. Results were impressive ( $r = .69$ ; Cicchetti, 1994) despite the fact that sample members had been receiving antidepressants. The follow-up EII scores also correlated significantly with self-report measures of adaptive functioning in the environment.

Also that same year, Adrian and Kaser-Boyd (1995) reported an examination of the construct validity of the EII by relating it to MMPI Ego-strength (Es) and Scale 8 scores in a heterogeneous group of adult psychiatric patients ( $n = 85$ ). As expected, overall EII scores were able to distinguish inpatients ( $n = 61$ ) from outpatients ( $n = 24$ ); however, only the *GHE* component of the EII was able to discriminate between psychotic ( $n = 37$ ) and non-psychotic individuals ( $n = 48$ ). None of the MMPI scales were able to distinguish any of the groups, but the EII and MMPI Scale 8 correlated positively ( $r = .25$ ;  $p < .05$ ) in the entire sample. Significant positive relationships were also found between the EII and several Scale 8 subscales, including Sc2 (Emotional Alienation), Sc5 (Lack of Ego Mastery, Defective Inhibition), Sc6 (Bizarre Sensory Experiences). More specifically, the



EII was significantly correlated with Sc5 and Sc6 in the nonpsychotic and inpatient groups, and the whole sample.

Perry, Sprock, et al. (1995) examined the discriminant validity of the EII by testing whether amphetamine affects cognitive functioning in a manner more consistent with anxiety or psychosis. Undergraduate males ( $n = 22$ ) were given amphetamine or a placebo, then administered the Rorschach and a measure of mood states. Three weeks later, all subjects experienced the counterbalanced condition. The authors hypothesized that amphetamine effects would mimic anxiety, not the disorganized and disinhibited thought processes in psychosis. The results confirmed this: EII scores did not differ between the amphetamine and placebo groups; of the CS scores, only those believed to assess stressful situations -- inanimate movement ( $m$ ) and shading variables ( $Y$ ) -- differed significantly.

To investigate the phenotypic similarities between adults with schizotypal personality disorder (STPD) and schizophrenia, Cadenhead, Perry, and Braff (1996) administered behavioral performance measures (e.g., VBM, critical stimulus duration [CSD]) and the EII to non-patient subjects ( $n = 21$ ) and individuals with STPD ( $n = 14$ ). They hypothesized that STPD subjects would demonstrate more impaired VBM and CSD performance, as well as higher EII scores, than non-patients. Results were encouraging but not unequivocal: Strong trends, but not significant differences, were found between the two groups. Still, there were significant associations between EII scores and VBM performance for the entire sample. Not surprisingly, STPD individuals who performed in the bottom 25% of the STPD group on the VBM task manifested a severity of thought disorder seen in schizophrenia.

Later, Auslander, Perry, and Jeste (2002) compared thought disorder severity among older individuals (aged 45 – 100) with paranoid ( $n = 27$ ) or non-paranoid schizophrenia ( $n = 17$ ) and a group of nonpatient controls ( $n = 45$ ). In keeping with the Perry et al. (1992) findings, the authors hypothesized that non-paranoid subjects would score significantly higher than paranoid subjects. Indeed, this was confirmed. Non-paranoid schizophrenia subjects scored significantly higher than paranoid subjects, who scored similar to the non-patient comparison group. When age, daily neuroleptic dose, and scores on the Dementia Rating Scale (DRS; Mattis, 1973), the Scale for the Assessment of Negative Symptoms (SANS; Andreason 1982), and the Adult North American Reading Test (ANART; Blair & Spreen, 1989) were used as covariates to examine group differences in EII scores, overall results did not change. Also, among the schizophrenia patients, EII scores correlated significantly with the DRS, but not with scores from the Scale for the Assessment of Positive Symptoms (SAPS; Andreason, 1984), Scale for the Assessment of Thought, Language, and Communication Disorders (TLC; Andreason, 1979), and the Hamilton Depression Rating Scale (HAM-D; Hamilton, 1967).

Perry et al. (2003) examined the EII's ability to distinguish among groups on the schizophrenia spectrum based on severity of disturbance in perception and thinking. The experimental groups included non-patients ( $n = 66$ ), college students with elevated scores on the MIS ( $n = 24$ ), 1<sup>st</sup> degree relatives of individuals with schizophrenia ( $n = 36$ ), individuals with STPD ( $n = 36$ ), outpatients with schizophrenia ( $n = 33$ ), and inpatients with schizophrenia ( $n = 56$ ). All participants were adults, and diagnostic groups were identified with the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders (4<sup>th</sup> ed. [DSM-IV]; SCID-I) and the SCID-II, for personality

disorders. As hypothesized, an analysis of variance (ANOVA) with focused linear contrasts demonstrated that EII scores increased successively across the groups. Concomitantly, the same focused linear contrast results were found for *FQ*-, *WSum6*, primitive contents, *PHE*, and *GHE*, but not *M*-. The authors concluded that the EII appears to be sophisticated enough to generate information about the specific severity levels of thought disturbances across the schizophrenia spectrum groups.

As noted earlier, Viglione, Perry, Jansak, et al. (2003) replaced the *HEV* with the *HRV* in Exner's (2003) CS. This was done because *PHE* scores tended to be more frequent, and occur with greater variability, than *GHE* scores in the literature. This trend was puzzling theoretically, and necessitated the use of a constant and weights in the original *HEV* equation to render *PHE* and *GHE* more similar statistically (re: mean, variability; Perry & Viglione, 1991). Viglione, Perry, Jansak, et al. addressed these problems by altering scoring criteria for human representational response content, COP, and several Special Scores. The resulting variables – *GHR* and *PHR* – were very similar statistically, which permits the calculation of *HRV* as the simple raw score difference between *GHR* and *PHR*.

Viglione, Perry, and Meyer (2003) then replaced the *HEV* in the EII with the *HRV* to create the EII-2. Rorschach data were collected from a large heterogenous sample of adults ( $n = 363$ ), and new factor coefficients were computed. Factor score coefficients for the EII and EII-2 are in Table 2.

Table 2

*EII / EII-2 Factor Score Coefficients*

CS Variable	EII Factor Score Coefficient	EII-2 Factor Score Coefficients
<i>Sum FQ-</i>	.136	.141
Primitive Contents	.068	.072
<i>GHE</i>	-.160	-.104
<i>PHE</i>	.108	.117
<i>M-</i>	.208	.198
<i>WSum6</i>	.050	.049
<i>R</i>	-.061	-.066
Constant	-.049	-.038

*Note.* EII = Ego Impairment Index; CS = Exner's (2003) Comprehensive System.

The EII and EII-2 were virtually indistinguishable ( $r = .99$ ). The authors also repeated the factor analytic procedures from the Perry and Viglione (1991) study, and again, found one robust factor which correlated  $r = .952$  with the original first factor. As a test of concurrent validity for the new index, the authors rank-ordered the clinical groups in their sample according to hypothesized severity of psychopathology. They concluded that the resulting correlation,  $r = .47$ , was close to the ceiling for this type of analysis given that the criteria (clinical subgroups) were not very precise.

Stokes et al. (2003) examined the predictive validity of the new EII-2 in a child inpatient psychiatric sample ( $n = 53$ , aged 5 – 12). The Rorschach was administered at intake, and the Devereux Scales of Mental Disorders (DSMD; Naglieri, LeBuffe, & Pfeiffer, 1994) was completed by parents at intake, 30 days after discharge (T2), and 120 days after discharge (T3). The DSMD is composed of three broad scales, each containing two factor-analytically derived subscales: Attention and Conduct subscales contribute to

the Externalizing Scale; Anxiety and Depression subscales make-up the Internalizing Scale; and Autism and Acute Problems subscales comprise the Critical Pathology Scale. The EII-2 was significantly related only to the Critical Pathology Scale at intake, and no significant relationships were found at T2; however, at T3, the EII-2 correlated significantly with all three DSMD Scales. The EII-2 predicted long-term, but not short-term, response to inpatient treatment, and did so better than severity of presenting symptomatology as measured by the DSMD. Interrater reliabilities (*ICC*) for the EII-2 variables ranged from .90 to 1.00.

In sum, the EII / EII-2 has demonstrated impressive interrater and test-retest reliability; it has demonstrated convergent and construct validity by relating in expected ways to self-report and behavioral / perceptual measures. Further, and impressively, it has discriminated among groups of individuals on the schizophrenia spectrum, and predicted response to psychotropic medication and inpatient psychiatric treatment. In a few studies, the *HEV*, or one of its variables, has contributed most to the EII's ability to distinguish groups. The validity data suggest that the EII-2 is sensitive to a core personality construct that is relatively independent of symptom presentation, and imperceptible to many other forms of assessment.

As noted, Viglione, Perry, and Meyer (2003) advocated using more behavioral markers of adaptive functioning as EII-2 criterion variables. Additionally, although the predictive and incremental validity studies of the EII-2 have been notable (Dawes, 1999; Perry, 2001; Perry & Viglione, 1991; Stokes et al., 2003), the scale has never been applied to long-term prospective study data, and it has never been used to predict a longitudinal health outcome such as mortality. Finally, a number of researchers have recommended

further investigation of the contribution of individual EII-2 components when the scale is used to predict an outcome via regression analysis (Perry & Viglione, 1991; Perry et al., 1992; Viglione, Perry, Jansak, et al., 2003).

#### *Long-Term Prospective Research and the Precursors Study*

In prospective research, the design of the study precedes baseline assessment, any intervention that might be used, and outcome assessment. Retrospective studies, in contrast, are designed after the phenomena of interest have occurred. In fact, retrospective designs typically rely on data collected for reasons other than the study itself. Long-term prospective designs include cohort studies, where a group of participants is followed for extended periods of time, e.g., years or even decades (Hess, 2004; Wadsworth et al, 2003). Such studies have been popular in the medical and social science literature because they are uniquely able to identify salient developmental phenomena, and risk factors of later functioning.

One of the earliest longitudinal studies, based on the Scottish Mental Survey, was initiated in 1932 to investigate the relationship between childhood IQ and mortality. All Scottish children born in 1921 and attending school on June 1<sup>st</sup>, 1932 ( $N = 87,498$ ), were tested with the Moray House Test (MHT), a measure of intelligence that has demonstrated substantial concurrent validity ( $r = .80$ ) with the Stanford-Binet. In subsequent decades, subsets of this cohort have provided health information to the organization responsible for the study, the Scottish Council for Research in Education. In a recent investigation, Whalley and Deary (2001) traced the members of the cohort who were living in Aberdeen, Scotland at the time of testing ( $N = 2,792$ ) in 1932. In the participants they were able to account for (80%;  $N = 2,230$ ), higher IQs at age 11 significantly reduced the

chance of death by age 76. More specifically, girls with IQs around 70 were about half as likely to survive to age 76 as were girls with IQs around 100. This long-term cohort study focused almost exclusively on intellectual functioning at baseline.

The Framingham Heart Study (FHS), in contrast, conducted a physical examination, and an interview about lifestyle details as baseline assessment. Initiated in 1948, the FHS was initiated in 1948 to identify risk factors for coronary heart disease in a mixed-gender cohort of 5,209. In 1971, 5,124 offspring (and their spouses) of the original cohort were enrolled (Amin et al., 2004), and the third generation is currently being recruited. Every two years since baseline assessment, participants have completed a physical examination, laboratory tests, and a medical history questionnaire (National Heart, Lung and Blood Institute [NHLBI], 2007). In this cohort, researchers have identified obesity as a risk factor for heart disease (Hubert, Feinleib, McNamara, & Castelli, 1983) and impaired cognitive performance (Elias, Elias, Sullivan, Wolf, & Agostino, 2005).

One of the largest prospective cohorts is from the Nurses' Health Study (NHS), initiated in 1976 by Dr. Frank Speizer to study long-term effects of oral contraceptive use in women. Over 122,000 married nurses between 30 and 55 years of age completed the baseline questionnaire, and participants still receive a biennial questionnaire packet inquiring about such topics as general health, smoking behavior, and hormone use. Questionnaires about diet were included in the 1980 packet, and have been included approximately every 4 years since then. Similarly, questions about quality-of-life were added in 1992, and included every 4 years. Toenail and blood samples have also been collected to identify potential nutrition and genetic markers (Laden et al., 2000). A second

Nurses' Health Study (NHS-II) cohort of about 117,000 participants was added in 1989 (Garland et al., 1999). Researchers have been prolific with these cohorts, e.g., there were approximately 100 publications on NHS and NHS-II data in 2005 alone. Among notable findings has been the association between higher intake of refined carbohydrates with increased risk of stroke (Oh et al., 2005).

In 1971, Moore, Gould and colleagues (2005) initiated a stratified, multistage, probability-sampling design to describe longitudinal patterns of alcohol use in this country. They created the National Health and Nutrition Examination Survey I (NHANES I), and administered it every year from 1971 through 1974. To incorporate data about health conditions and mortality, the NHANES Epidemiological Follow-up Survey (NHEFS) was also administered several times between 1982 and 1992. The sample consisted of more than 20,000 non-institutionalized, non-military U.S. citizens between the ages of 1 and 75. With about 14,000 of these participants, Moore et al. (2005) found that overall alcohol use decreased with age, and higher consumption was associated with being a white, married male, having more education and a higher income, being employed, and having a smoking habit. The following year, Moore et al. (2006) focused on the mortality risks of alcohol use among older adults ( $N = 4,691$ ), and found that for men only, drinking "risky" amounts of alcohol was associated with greater mortality (hazard ratio = 1.20).

One of the first cohort studies to utilize standardized psychological measures was the Normative Aging Study (NAS). Established in 1963 by the Veteran's Administration, it was designed to investigate the relationships between the aging process and the natural history of chronic diseases (Bell, Rose, & Damon, 1972). It consists of a cohort of over



2,200 socio-economically diverse men who ranged in age from 21 to 80 at the time of initial assessment. Volunteers undergo medical exams every 3 to 5 years, and routinely provide information about general health habits, food intake, and emotional and cognitive functioning. In one study, Kawachi, Sparrow, Spiro, Vokonas, and Weiss (1996) utilized the MMPI-2 Anger Content Scale to identify subjects who reported high versus low amounts of anger. The former group was 3 times more likely to experience nonfatal myocardial infarction (MI) than the latter.

Taking an exclusively psychological focus, Block and colleagues (Block, Block, & Keyes, 1980; Shedler & Block, 1990) initiated a longitudinal study of ego and cognitive development in 1975 with 130 children recruited from a university or parent-cooperative preschool. Baseline assessment of a variety of psychological measures began when participants were 3 years old, and follow-up assessments occurred at 4, 5, 7, 11, 14, 18, 23, and 32 (Onishi, Gjerde, & Block, 2001). The participants are heterogenous regarding ethnicity, socio-economic class, and parents' education. In one study, Cramer and Block (1998) examined Vaillant's concept of a defense mechanism hierarchy by relating personality characteristics at age 4 to type of defenses used at age 23. Males tended to demonstrate continuity between preschool and adult defense mechanism use, whereas females did not. For males, emotional immaturity and poor social functioning predicted later use of denial. Females who used denial as an adult tended to withdraw into fantasy as children.

One of the longest-running long-term cohort studies has been the Study of Adult Development at the Harvard University Health Service. Data collection on Harvard-educated and inner-city male cohorts began around 1940 ( $N = 724$ ) with the purpose of

investigating adaptation to stress and the use of defense mechanisms, the effects of affective disorders on health and mortality, childhood risk factors of poor adaptation later in life, the natural history of alcoholism and drug abuse as diseases. Concomitantly, a number of physical and psychosocial health variables have been assessed over the years, including body mass index, years of education, maturity of psychological defenses, warmth of childhood, and physical and mental health status, among many others (Cromie, 2001). Dr. George Vaillant, a recent director of the study, has published widely with this data (Vaillant, 1977, 1983, 1995, 2002). In one study, for example, tranquilizer use before the age of 50 strongly predicted poor physical and mental health outcomes after age 65 ( $n = 173$ ); concomitantly, maturity of psychological defenses predicted positive health outcomes (Vaillant & Vaillant, 1990). In a later study, a relative lack of nicotine and alcohol use before age 50 appeared to serve a protective function in the healthy aging process. Marriage stability was also a significant predictor of subjective life satisfaction and objective mental health functioning ( $n = 569$ ).

Rivaling the length of the Adult Development Study is the Johns Hopkins Precursors Study (JHPS), started by Dr. Caroline Bedell Thomas in 1946 to investigate risk factors (in young adulthood) of health problems and death in later life. Like the Framingham Heart Study, the NHS/NHS-II studies, the NHANES, and the NAS, the JHPS included a variety of physical and physiological measures, and questionnaires about medical history, diet, and health practices during baseline and follow-up assessments. Further, like the Block studies and the Study of Adult Development, the JHPS also collected an abundance of baseline and follow-up data from psychosocial and

psychological measures. However, it is unique among these long-term cohort studies in that it contains Rorschach data.

As noted, the JHPS cohort consists of 17 Johns Hopkins medical school classes which graduated from 1948 through 1964 ( $N = 1,337$ ; Thomas, 1976). At baseline, data on approximately 6,000 medical, psychological, and psychosocial variables were collected (L. Meoni, personal communication, Sept. 14<sup>th</sup>, 2007), and most of the surviving participants still complete annual follow-up questionnaires (Graves, Mead, Wang, Liang, & Klag, 1994; Torre et al., 2005). Psychological risk factors for later negative health outcomes, including mortality, have been a popular focus of Precursors researchers over the years. Thomas (1976), for example, reported a significant association between premature diseases (including those that are fatal) and early levels of anger, anxiety, insomnia, and smoking and drinking behavior.

Rorschach tests were administered on an individual basis ( $n = 453$ ), or in group format ( $n = 484$ ) at baseline. Individual administrations were conducted according to Beck's guidelines (Beck, 1944), while group administrations utilized procedures adapted from Beck by H. Barry Molish and Ellen S. Freed (Thomas, Ross, Brown, & Duszynski, 1973), as well as a group administration form described by Harrower and Steiner (1951; Shaffer, Duszynski, & Thomas, 1981). Rorschach protocols were originally scored with Beck's system as well, and this data has been used in a number of studies.

In one of the earliest published JHPS studies using the Rorschach data, Molish, Molish, and Thomas (1950) hypothesized, and found, that "superior" individuals demonstrate higher than average response productivity. Individually-administered protocols for the entire Johns Hopkins Medical School class of 1948 ( $n = 60$ ) contained a

total of 3,300 responses (10 cards), with an average  $R$  of 55.37 and a standard deviation of 24. For this group,  $R$  ranged from 18 to 124.

Thomas and Duszynski (1985) searched for the word “whirling,” and some of its synonyms (collectively, whirlall words), in the Rorschach records of all cohort members who took that test at baseline ( $n = 1154$ ). Interestingly, members in the suicide and other deaths (accidental, medical issues) groups had used a larger proportion of whirlall words than members in healthy or disease groups. Further, deceased cohort members had used whirlall words 3 times more frequently than ones who were still alive.

Graves and Thomas (1981) created the Rorschach Interaction Scale (RIS) to assess the themes of emotional relatedness in the free response phase of Rorschach responses. RIS scores allowed cohort members ( $n = 319$ ) to be classified as Flexible, Conformist, Emphatic Positive, Controlled, Ambivalent, or Avoidant in their habitual style of interaction. These styles were subsequently related to midlife health status groups (Healthy, Cardiovascular, Mental Disorder, Cancer). Results indicated that participants who later developed cancer and mental health problems tended to rely on Ambivalent (emotional lability) and Controlled (restricted emotional involvement) interaction styles, at least at baseline. In a similar study, Graves, Mead, and Pearson (1986) found that Controlled, Ambivalent, and Avoidant participants were two, three, and four times as likely to develop cancer respectively than Flexible participants after 30 years.

Thomas (1976) examined the relationships between the baseline measure of habitual stress responses (HNT) and later occurrence of suicide, mental illness, malignant tumors, coronary occlusion, and hypertension. JHPS Cohort members who developed any of these diseases scored significantly higher on the HNT than did members who remained

healthy. Differences among the disease groups emerged as well. The mental illness group reported more depression, anxiety, anger, and insomnia at baseline than controls, while the coronary occlusion group reported more depression, anxiety, and tiredness upon waking. The malignant tumor group scored like, or below, controls on depression, anxiety, and anger, but they reported more alcohol use. The total disorders group smoked cigarettes more frequently than did the control group. In the same study, Thomas found that the suicide, mental illness, and malignant tumor groups obtained significantly lower mean scores on the Closeness to Parents Scale than did other groups. This scale measures perceived subject-parent relationship quality.

Shaffer, Duszynski, and Thomas (1982) revisited the Closeness to Parents Scale, but focused exclusively on the incidence of cancer in the JHPS cohort. Individuals who had developed cancer ( $n = 25$ ) reported significantly poorer relationships with their fathers at baseline than did individuals who were still free of major illnesses like heart disease, cancer, and mental illness ( $n = 318$ ). When risk factors for cancer, such as alcohol and nicotine use, weight, serum cholesterol level, radiation exposure, and geographic area, were controlled-for, differences between the two groups remained significant.

Chang et al. (1997) examined the relationship between sleep habits during medical school and the subsequent incidence of depressive episodes. Cohort members who simply answered “yes” to the question “Do you ever have insomnia?” had a significantly higher cumulative incidence of depression in midlife than individuals who answered “no.” Further, when the possible influence of age at graduation, class year, parental history of depression, temperament, and coffee intake were controlled for, cohort members who reported insomnia were still twice as likely to develop depression than those who did not.

A similar, but less pronounced, pattern of results was found for those who reported sleep difficulties when stressed, or poor quality of sleep. Results were strengthened by the fact that cohort members who had depression at baseline were excluded from the analyses.

Graves et al. (1994) used latent class analysis to identify three temperament types (Tension-in, Tension-out, Stable) based on responses to eight items from the JHPS Habit Survey Questionnaire (HSQ). Tension-in individuals are anxious and inhibiting, and express tension through psychophysiological pathways. Tension-out individuals are volatile, and express tension through restlessness and emotional expressiveness. Stable individuals are described as “solid, stable, self-contained” (Graves et al., 1994, p. 118). Graves et al. found that the cumulative incidence of mortality for Tension-in cohort members was significantly higher than for Tension-out or Stable members. When data were controlled for age at medical school graduation, serum cholesterol level, nicotine use, and a weight to height ratio (Quetelet Index), Tension-in individuals were more than 1.5 times as likely to die during the entire follow-up period than were Stable individuals. Tension-out and Stable individuals did not differ significantly. When data were restricted to only include individuals aged 55 and younger, Tension-in individuals were more than 2.5 times as likely to die than the Stable group.

A number of the variables collected from the HPS cohort at baseline and follow-up relate to adaptive functioning, the construct targeted by the EII-2. For example, mortality, incidence of divorce (Betz & Thomas, 1978), incidence of depression (Chang et al., 1997), and alcohol use (Thomas, Santora, & Shaffer, 1979) are all relevant, and can facilitated examinations of the construct validity of the EII-2.

## *Hypotheses*

The proposed study utilized the JHPS data. Hypotheses were as follows:

1. Given the high rate of comorbidity between depression and personality pathology (Hirschfeld, 1999), it was hypothesized that EII-2 scores would significantly predict the incidence of depression over time until age 57. This age was chosen because it represents a natural cut-point in the bimodal distribution of ages at which male cohort members experienced their first depressive episode; and this age is consistent with epidemiological data that suggests cerebrovascular disease is a major contributor to depression onset after age 50 or 60 (Alexopoulos, Meyers, Young, Campbell, Silbersweig, & Carlson, 1997; Van den berg, Oldehinkel, Bouhus, Brilman, Beekman, & Ormel, 2001).

2. Extending on the hypothesized relationship between ego impairment and depression, individuals with more ego impairment as a group should have more depression than individuals with healthier ego functioning. Therefore, High EII-2 scorers (top 50% of cohort) would have a significantly greater cumulative incidence of depression over time (until age 57) than low EII-2 scorers (bottom 50%), and the difference would be significant.

3. Since the EII-2 appears to measure a trait-level personality construct related to coping and adaptive functioning (Perry & Viglione, 1991), the overall EII-2 score would significantly predict later life psychological health.

4. Consistent with recommendations from Perry and colleagues (Perry & Viglione, 1991; Perry et al., 1992; Viglione, Perry, Jansak, et al., 2003), the relative contributions of the individual EII-2 components were examined. Previous data suggest that the primary

strength of the EII-2 is in the interpersonal variables (*GHR, PHR*; Perry & Braff, 1994; Perry & Viglione, 1991), followed by *WSum6* (Perry et al., 2003), *M-* (Adrian & Kaser-Boyd, 1995), and finally, primitive contents. It was hypothesized that the interpersonal components of the EII-2 (*GHR, PHR*) would predict later life psychological health better than the thought disorder component (*WSum6*), which would predict better than a combined interpersonal / thought disorder component (*M-*); and all of these components would predict better than primitive contents component, which has never been found to be a significant EII / EII-2 component in prediction.

5. Psychiatric illness appears to increase the risk of mortality in general (Drew, 2005). More specifically, depression is an independent risk factor for hypertension, and nicotine dependence is higher among the psychiatric disorders (Yates & Brooks, 2001). Further, suicides, homicides, and deaths from unnatural causes occur more frequently among the mentally-ill (Hiroeh, Appleby, Mortensen, & Dunn, 2001). Therefore, it was hypothesized that EII-2 scores would predict the incidence of mortality over time.

6. Further, it was hypothesized that Kaplan-Meier survival curves of the incidence of mortality over time would be significantly steeper (more cumulative incidence) for high EII-2 scores than low EII-2 scorers, and the difference in incidence would be significant.

If reality-testing is relatively intact (*Sum FQ-*), thought processes are relatively free of disruption (*WSum6, M-*), human beings are perceived realistically and logically (*GHR, PHR, M-*), and defensive processes effectively prevent primitive thought content from intruding into conscious awareness (Primitive Contents), then adaptive functioning will be good, and internal resources for coping with demands of both internal and external



stressors will be more available (low EII-2 score). Theoretically, individuals with good coping and adaptive capacity are (a) less likely to use drugs or alcohol to cope with internal or external stressors; (b) more likely to muster the foresight, discipline, effective problem-solving, and interpersonal effectiveness required over time to find, or construct, fulfilling job environments; (c) more likely to perceive complex situations accurately, less likely to be influenced by unconscious thoughts, wishes, and fantasies; (d) less likely to have medical problems over time because they are more likely to muster the foresight and discipline to maintain a long-term personal health program, i.e., take care of themselves better; (e) more likely to be interpersonally effective, perceive others realistically, and less likely to be influenced by unconscious thoughts, wishes, and fantasies, and so, more likely to develop and maintain satisfying social support networks over time; and (f) more likely to enter marriage with realistic perceptions of their partners, and so, less likely to divorce. Therefore it was hypothesized that:

7. EII-2 scores would significantly predict the cumulative incidence of divorce over time.

8. EII-2 scores would correlate significantly and positively with alcohol use (higher EII-2 scores reflect more ego impairment).

9. EII-2 scores would correlate significantly and positively with job dissatisfaction.

10. EII-2 scores would correlate significantly and negatively with perceived social support.

11. EII-2 scores would correlate significantly and negatively with the physical health over time.

12. EII-2 scores would correlate significantly and negatively with the use of preventive health services.

13. Finally, it was hypothesized that an aggregate variable based on all of these variables (but mortality) would correlate significantly and positively with the EII-2.

## Chapter 2

### Method

#### *Participants*

The participants in this study were drawn from the Johns Hopkins Precursors Study (JHPS) longitudinal cohort. The JHPS was initiated by Dr. Caroline Bedell Thomas in 1947 in order to investigate youthful ‘precursors’ to later disease and death. The cohort consists of the 17 Johns Hopkins Medical School classes that graduated from 1948 through 1964, a total of 1,337 individuals (Thomas, 1976). As students who have been accepted into a prestigious medical school, cohort participants have higher IQs than the normal population, and the range of IQs is restricted. Molish, Molish, and Thomas (1950), for example, found that the average Wechsler-Bellevue Full Scale IQ of the class of 1948 was 131, with a standard deviation of 6.45. Likewise, the socioeconomic class of the cohort participants is higher than average, with a restricted range.

Baseline and follow-up assessment for most JHPS participants included a wide variety of physiological, psychological, and psychosocial variables, including height, weight, blood pressure, heart rate, total serum cholesterol level, eating and smoking habits, sleep habits, coffee and alcohol intake, results from specific medical tests (e.g., electrocardiogram, urinalysis), medical history of parents, interpersonal attitudes, marital status, and information on religious and educational background to name just a very few. Some data were gathered only at baseline, including a vocational interest survey, a figure

drawing test, and the Rorschach inkblot test. Since graduation, participants have been followed yearly with questionnaire mailings, and more comprehensive surveys every 5<sup>th</sup> year; data collection continues (Wang et al., in press). Response rates for any 5-year period have been high, ranging from 87 to 94% (Chang et al., 1997), and the participants' self-report measures regarding health and disease have been found to be extremely accurate (Klag et al., 1993).

A vast majority of the cohort completed the Rorschach test at baseline ( $n = 937$ ); 453 of these were individual administrations according to Beck's (1944) guidelines, while 484 were group administrations using procedures adapted from Beck by H. Barry Molish and Ellen S. Freed (Thomas, Ross, Brown, & Duszynski, 1973), and a group administration form described by Harrower and Steiner (1951; Shaffer et al., 1981). Only data from white males were utilized in this study because they comprise approximately 90% of the cohort. Further, only those participants whose Rorschach records contain between 18 and 50 responses were utilized. This resulted in a sample size of 412. Dean et al. (2007) demonstrated that constraining  $R$  to the range of 18 to 40 did not reduce the validity of CS scoring.

During the era when baseline assessment occurred (1948 – 1964), obtaining informed consent from participants was not typical. However, after the Johns Hopkins University Joint Committee on Clinical Investigation was established, it approved the JHPS follow-up procedures (Chang et al., 2002). The University of Toledo's equivalent body, the Institutional Review Board (IRB), approved the current study on June 5<sup>th</sup>, 2008 (IRB #106084).

## *Measures*

*The Rorschach EII-2.* Hermann Rorschach's inkblot method was published in 1921, but did not receive an English translation until 20 years later (Rorschach, 1921/1942). A number of systems for administering Rorschach inkblots, and scoring responses to them, have been described since the test's inception, but Exner's (2003) CS is the most psychometrically sound, and the most commonly used and taught (Anastasi & Urbina, 1997; Mihura & Weinle, 2002). It was used in the current study because scoring for the EII-2 is derived from CS variable scores. However, the Rorschach was originally administered to JHPS cohort members using Beck's guidelines (Beck, 1944; Beck et al., 1961). Like CS administration guidelines, they include response and inquiry phases where examinees provide, respectively, a basic response, then elaboration of the response. The JHPS response material was recorded into a computer, then translated into an Excel file. Inquiry material was not so translated, but the original administration forms, including inquiry material, were electronically scanned into PDF documents. It was important for the coders in this project to have access to the inquiry material because it is particularly rich, and many CS variables require this extended sampling of behavior to be accurately scored (Exner, 2003). This is especially true of the Special Scores, which are sensitive to thought disorder. Subsequently, CS variable values were used to calculate EII-2 scores for each participant (Please see Introduction for a discussion about reliability of the EII/EII-2).

*Health Behavior Questionnaire (HBQ).* The HBQ is a lengthy self-report measure that assesses alcohol, caffeine, and nicotine use, amount and type of daily or weekly physical activity, appetite, and diet, including food types and typical portion sizes.

Portions of the HBQ have been used to demonstrate associations between coffee intake and blood pressure (Klag et al., 2002), coffee intake and coronary heart disease (LaCroix, Mead, Liang, Thomas, & Pearson, 1986), and midlife alcohol intake and coronary heart disease (Thomas, Santora, & Shaffer, 1980). In the current study, only the portion on alcohol use was used. It consists of five yes-no items about excessive use, four open-ended questions about amount and type of alcohol ingested on a weekly basis, and one multiple-choice item on the absolute frequency of alcohol use. The HBQ was administered in follow-up packets on the following years: 1978, 1986, 1989, 1993, 1997, and 2003. No reliability data have been reported on this measure. Individuals tend to underestimate or purposely underreport the amount of alcohol they consume on self-report measures (Fuller, Lee, & Gordis, 1988), though Sommers, Dyehouse, Howe et al. (2000) found that men were more accurate overall than women. The JHPS cohort members tend to be accurate self-reporters (Klag et al., 1993).

*Family and Career Questionnaire (FCQ).* The FCQ is a self-report measure that assesses the number, type, and quality of social relationships, and work environment and satisfaction. It was administered in the follow-up packets of 1987, 1991, and 1996. In the current study, only specific items pertaining to job dissatisfaction, divorce, and perceived social support were used. The four job dissatisfaction items were adapted directly from the Job Content Questionnaire (JCQ; Karasek, 1985), which is a popular measure of psychosocial work environment in North America and Europe (Johnson, et al., 1995). Lichtenstein (1984) used an earlier version of this scale, the Quality of Employment Survey (Quinn & Staines, 1977) to predict physicians' intention to quit jobs in prison health settings. More recently, Johnson et al. (1995) demonstrated that more job

dissatisfaction, as measured with the FCQ, was associated with higher job demands. Internal consistency reliability of the four job dissatisfaction items specifically was .79 (Cronbach's alpha). An example of a one of these items is "On a scale from 1 (extremely satisfied) to 7 (extremely dissatisfied), all in all, how satisfied are you with your work?"

The first four items of the FCQ administered in 1987 were yes-no and open-ended items about marital status, number of marriages in the past, and how up to three marriages ended, e.g., divorce, separation, death of spouse. Mead, Wang, and Klag (1997) used these items to show that the cumulative incidence of divorce in this cohort was significantly related to chosen medical specialty. Reliability data have not been reported for this measure.

*Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36)*. The SF-36 was created by Ware and Sherbourne (1992) to assess general health status. Subscales were designed to assess the following: Physical functioning, role limitations because of physical problems, social functioning, physical pain, general mental health, role limitations because of emotional problems, vitality, and general health perceptions. Individual items have three, five, or six anchor points. An example of a six point items is "On a scale of 1 (All the time) to 6 (None of the time), How much time in the past four weeks have you been downhearted and blue?" The SF-36 is used regularly in the medical treatment outcome literature. It has good internal consistency, and construct validity (McHorney, 1996; Ware & Sherbourne, 1991). Stewart, Greenfield, Hays, Wells, Rogers, Berry, et al., (1989), for example, obtained good Cronbach's alpha coefficients for the mental health (Cronbach's alpha = .88) and physical functioning (alpha = .86) scales which were used as criterion variables in the current study. The Physical Health subscale

has effectively distinguished between patients with minor and major medical illnesses, while the Psychological Health subscale has been able to distinguish between (a) psychiatric patients and individuals with serious medical illnesses, and (b) psychiatric patients who differed by symptom severity (McHorney, Ware, & Waczek, 1993). The SF-36 was administered in the 1992, 1995, 1998, 2001, and 2005 follow-up packets. Data from all administrations were used in this study.

*Preventive Health Services Questionnaire (PHSQ).* The PHSQ is a self-report measure with 10 yes-no and multiple-choice items that assess participants' willingness to obtain medical tests, e.g., "Have you seen a physician in the past two years to assess your risk for future disease?" It was administered only in the 1997 follow-up packet. Gross, Mead, Ford, and Klag (2000) found that cohort members with no regular source of medical care in 1991, as assessed by the Medical Care Questionnaire in the 1991 packet, tended to have lower PHSQ scores (fewer preventive services) in 1997. Reliability data have not been reported on this measure.

*Social Support and Network Questionnaire (SSNQ).* The SSNQ is a self-report measure with 19 Likert items about perceived quality of social support, three multiple-choice items about marital status and frequency of contact with a closest friend, and 10 open-ended items assessing the numbers of friends and relatives. It was only administered in 2002, and has not been used in a research study to date.

*Depression.* The cumulative incidence of depression over time was used as the criterion variable in one of the Cox regressions and one of the Kaplan-Meier analyses. A committee of five physician reviewers has retroactively assigned an International Classification of Diseases, 9<sup>th</sup> Revision (ICD-9) diagnosis of depression whenever a



participant self-reported an episode of depression in the follow-up questionnaires, or when a physician's note indicated that the participant experienced such an episode. Episodes of depression were excluded if they were related to grief, or if they lasted less than two weeks (Chang et al., 1997; L. Meoni, personal communication, Sept. 14<sup>th</sup>, 2007). Data have not been reported regarding the reliability or accuracy of diagnostic judgments involved in constructing this variable.

*Mortality.* The cumulative incidence of mortality over time was used as a criterion variable in one of the Cox regressions and one of the Kaplan-Meier analyses. Vital status of participants and cause of death (including suicide) have been determined by searching the National Death Index (NDI), obtaining death certificates, searching obituaries, and calling family members, friends, and classmates.

*Covariates.* Originally, documented psychological risk factors for depression and mortality in this cohort were going to be used as covariates in the Cox regression analyses. Baseline risk factor data were to be obtained on insomnia, temperament, closeness to parents, and habitual responses to stress; however, in the interest of retaining as much power in the analyses as possible, risk factor covariates were dropped from the primary Cox and multiple linear regressions. Secondary Cox and multiple linear regressions were to be conducted only in the event that robust effects were found

#### *Procedures*

JHPS variable data were organized and screened for data entry error by generating a table with minimum and maximum scores, standard deviations, means, medians, modes, and values for kurtosis and skew. Univariate and multivariate diagnostics were conducted to determine if extreme outliers, multicollinearity, or heteroscedasticity were

present in the data. Further, diagnostics were completed to determine if the assumption of proportional hazards was met for the survival analyses.

To ensure coding accuracy, all coders were required to meet minimum accuracy benchmark standards for single category CS variables (Pairs, Popular), multi-category variables (DQ, FQ, Z), and multivariate coding segments (Location & Space, Determinants, Contents, Special Scores) before they could begin coding the JHPS protocols. Scored protocols published by Exner and Erdberg (2005) were chosen as the “gold standard” agreement benchmarks. Agreement was calculated using Janson and Olsson’s (2001) Iota statistic, which is a measure of chance-corrected agreement that can be used when (a) measures are multivariate, (b) agreement is computed at the response or summary-score level, and (c) two or more coders have contributed agreement data. In this study, agreement was computed at the summary-score level. Iota magnitudes are interpreted much like *kappa* or *ICC* magnitudes. In the current study, minimum benchmark standards for FQ and Special Scores were Iota = .60 because these variables are consistently more challenging to code (Meyer, Erdberg, & Shaffer, 2007). The minimum benchmark for the remaining variables was Iota = .80.

All Rorschach scoring was conducted blind to other variables in this study. To determine interrater reliability for the CS coding of the JHPS protocols, 36 of the study author’s batch of protocols were randomly selected, then distributed among six of the other independent coders. One-way, random effects, absolute agreement intraclass correlations (*ICC* [1]; McGraw & Wong, 1996) were computed on coding judgments between the study author and one other coder. Two of these independent coders (JM, GM) are Associate Professors in Clinical Psychology at the University of Toledo with

extensive experience in administering, coding, researching, and teaching the Rorschach test. The remaining coders (ND, SH, NK, EK, AU) are students in the clinical psychology program at the same university who have completed a personality assessment course, and have extensive experience coding with the CS.

After coders met accuracy standards, individually-administered Rorschach protocols ( $R = 18 - 50$ ) from white male cohort members were scored with the EII-2 / CS variables, and resulting variable values were entered into an SPSS database. SPSS syntax was used to obtain final EII-2 scores for all participants; the equation, based on factor score coefficients in Viglione et al. (2003), was as follows:  $EII-2 = .141 (Sum FQ-) + .049 (modified WSum6) + .072 (Primitive Contents) + .198 (M-) + .117 (PHR) - .104 (GHR) - .066 (R) - .038$ .

Finally, the reliability of each self-report measure used in this study was examined using Cronbach's alpha coefficient (Cronbach, 1951). Alpha assesses internal consistency of the items in a scale. Higher scores indicate that scale items are measuring a unidimensional construct.

### *Analyses*

All statistical analyses were completed with the SPSS (2007), version 16.0. To examine hypotheses regarding the predictive validity of the EII-2, Cox regressions and Kaplan-Meier analyses were conducted. Both techniques treat time as the DV. The Cox regression examines the effects of an IV and other covariates on time to a criterion event, (e.g., death, first depressive episode, first divorce). The technique generates a regression equation that predicts survival time. It also generates a hazard ratio, or odds ratio, for each covariate, where higher values (above 1.0) indicate that the covariate increases the

probability of the occurrence of the event, and lower values (below 1.0) decrease the probability of the occurrence of the event (Tabachnick & Fidell, 2001). Significance levels for entry and removal of variables at each step of the analysis were, respectively,  $p < .15$  and  $.20$ . All Cox regressions used final EII-2 scores as the IV. DVs for the three Cox regressions were time to first depressive episode, time to death, and time to first divorce.

The Kaplan-Meier analysis calculates the cumulative incidence of a specified event (e.g., death, or first episode of depression) over time. This information can be presented graphically in plot form. A log-rank test can be used to determine if the resulting survival curves, or functions, of two groups differ significantly (Tabachnick & Fiddell, 2001). Covariates are not used. In the first Kaplan-Meier analysis, functions for the DV (incidence of depression over time before age 57) were calculated for each of two groups – high and low EII-2 scorers – then the log-rank test was calculated to determine if the survival functions for these groups differed significantly. Next, functions for the cumulative incidence of mortality over time were calculated for the two groups, and the log rank test was calculated to determine if the survival functions for the groups differed significantly.

A linear regression was conducted to examine how well the overall EII-2 score predicted later life psychological health, as measured by the SF-36. Another linear regression was conducted to investigate the contribution of individual EII-2 components. Primitive contents were entered as block 1, *M-* was entered as block 2, *WSum6* was entered as block 3, and the interpersonal variables (*GHR*, *PHR*) were entered as block 4. Later life psychological health was the DV. As with the Cox regression, significance

levels for entry and removal of variables at each step of the analysis were, respectively,  $p < .15$  and  $.20$ .

#### *Calculation of Self-Report and Aggregate Dependent Variables*

*SF-36 psychological health.* The SF-36 was administered in 1992, 1995, 1998, 2001, and 2005. Scores on the psychological health scale were recoded in the database so that higher values indicated more health. Subsequently, scores were converted to Z-scores for each administration, then the mean of these Z-scores was computed as the final score for analyses. An example of one of the five SF-36 psychological health items (item 9 *b, c, d, f, h*) is “On a scale of 1 (All the time) to 6 (None of the time), How much time in the past four weeks have you been downhearted and blue?” Final scores were expected to correlate negatively with the EII-2.

*Number of divorces.* Item 2 in Section I (Social Relationships) of the 1987 FCQ (described above) assesses age at divorce (up to three total), and number of years that each marriage lasted. EII-2 scores were the IV, and time to first divorce was the DV. The overall EII-2 scores were expected to predict the cumulative incidence of divorce over time.

*Alcohol use variable.* The alcohol use items from Section V (items *a* through *j*) of the HBQ administered in 1978, 1986, 1989, 1993, 1997, 2003. Scores for each of the ten items were summed for each administration, then averaged across the seven administrations for each participant. Higher values indicate more alcohol consumption, so it was hypothesized that higher scores would correlate positively with the EII-2.

*Job dissatisfaction variable.* The four job dissatisfaction items adapted from the JCQ (Karasek, 1985) were included as part of the lengthier Job Satisfaction

Questionnaire administered in 1987, 1991, and 1996. For each item, the mean score across administrations was obtained, then converted to a Z-score. The mean of these Z-scores across the four items was computed as the final job dissatisfaction score. Higher values indicate more job dissatisfaction, so they were expected to correlate positively with the EII-2.

*Social support variable.* The mean of the scores on SSNQ items 31 – 49 (1994 and 2002 follow-up packets) was obtained for each administration, then the mean value across the two administrations was computed as the final SSNQ score for each participant. Higher values indicate that social support is perceived as more available. An example of these SSNQ items is “On a scale of 1 (Never) to 5 (Always), How often is someone available to you who understands your problems?” The SSNQ was expected to correlate negatively with the EII-2.

*SF-36 physical health variable.* As part of the larger SF-36, this questionnaire was administered in 1992, 1995, 1998, 2001, and 2005. Scores on the SF-36 physical health scale were converted to Z-scores for each administration, then the mean of these Z-scores was computed as the final score for analyses. An example of one of the ten SF-36 physical health items (item 3 *a* through *j* in follow-up packets) is “On a scale of 1 (Limited a lot) to 3 (Not limited at all), How much does your physical health limit bathing or dressing yourself?” Higher values indicate less physical health concerns during the assessment period; scores were expected to correlate negatively with the EII-2.

*Preventive health variable.* The PHSQ was administered in 1997. The first eight items pertain to having obtained different types of preventive medical screening tests in the past five years. They were summed to obtain a final PHSQ Prevention Index score for

each participant. High values indicate more use of preventive services, so the Prevention Index was expected to correlate negatively with the EII-2.

Self-report variables were calculated from data that were available whenever cases did not have data for a specific self-report measure for any year. For example, if a case had data from three of four administrations of a specific questionnaire, this variable score was calculated (averaged) using the three administrations. When cases completed only some of the questionnaire items for a specific administration, a minimum number of completed items was required for data from that administration to be included in the analysis. Social support variable data, for instance, was only included for a specific year if at least 10 of the 18 items were completed. When cases had no data whatsoever for a self-report variable, they were excluded from the analysis.

*Total Impairment aggregate variable.* The aggregate variable was constructed from the self-report variables described above, as well as the dichotomous variables indicating if a cohort member has, or has not, been depressed (before age 58) or divorced (before 1988). The impairment variable was computed by summing Z-scores for the depression, divorce, alcohol use, job dissatisfaction, social support, physical health, psychological health, and preventive health variables, but only for cases who had data for all eight of the variables. The Z-scores for the social support, preventive health, psychological health, and physical health variables were multiplied by -1 to change their sign because they were expected to correlate negatively to the EII-2. This variable was expected to correlate positively with the EII-2.

*Moderator Variables and Secondary Analyses.*

Coders in this study made coding decisions based on response material from an Excel file, and inquiry material directly from electronically-scanned (PDFs) copies of the original administration forms. During this process, it became apparent that there was a great deal of variability among the eight original examiners in how response and inquiry material was manually recorded (written) during the actual administrations. Some recorded a brief sentence as the response, then wrote percept locations in the inquiry column of the record form. Others required extra record form pages because they recorded so much material. An exploration of the possible effects of such examiner differences was not part of the proposed study; however, it was included in a set of secondary analyses conducted after the primary analyses from the proposal were finished. This involved re-running the linear and Cox regressions from the primary analyses using examiner variables as an initial step, entered before the EII-2. Further, to examine the possible influence of examiner differences on the relationships between the EII-2 and the self-report construct validity variables, the original bivariate correlations (hypotheses 8 – 13) were replaced with linear regressions, where examiner variables were entered as the initial step. Examiner variables were simply constructed by giving each examiner a value of 1.

Another moderating variable was also included in the secondary analyses. Rorschach protocol complexity indicates the degree to which cognitive flexibility, motivation, and problem-solving capacity are present in the protocol (Dean et al., 2007). Some data suggest that protocol complexity can moderate the Rorschach's predictive power (Morgan & Viglione, 1992). Dean et al. (2007), for example, found that complexity enhanced the Perceptual Thinking Index's (Exner, 2003) ability to predict thought



disorder scores on a self-report measure. To examine the extent to which protocol complexity influences the predictive ability of the EII-2, it was entered in a separate step after the examiner variables, and before the EII-2 or its components, in the Cox and linear regressions. The results from these secondary analyses were reported at the end of the Results section.

## Chapter 3

### Results

#### *Interrater and Scale Reliability*

All Rorschach protocols were originally administered and scored between 1948 and 1964 using Beck's guidelines (Beck, 1944; Beck et al., 1961). Individually-administered protocols ( $n = 453$ ) were re-coded using Exner's (2003) CS to generate components for the EII-2 score. The batch of cases for final analyses was limited to white males ( $n = 412$ ); the study author coded 227 of these, and the remaining 185 protocols were coded by seven different coders. Coder accuracy and interrater reliability procedures were described earlier.

Results of the interrater reliability analyses are presented in Table 3. Agreement was in the excellent range (Cicchetti, 1994) for all but two of the study variables. Agreement was good for *Sum FQ-* and fair for *M-*. Importantly, agreement on the *EII-2* total score was excellent (Cicchetti, 1994).

Table 3

*Interrater Reliability Coefficients for EII-2*

*and Components*

Variable / Score	ICC	95% CI
<i>Sum FQ-</i>	.69**	.48 - .83
Primitive Contents	.94**	.89 - .97
<i>GHR</i>	.83**	.70 - .91
<i>PHR</i>	.91**	.83 - .95
<i>M-</i>	.56**	.29 - .75
<i>WSum6</i>	.80**	.65 - .89
<i>R</i>	1.00**	-
<i>EII-2</i>	.87**	.76 - .93

*Note.*  $n = 36$  for all variables; CI refers to Confidence Interval.

\*\*  $p < .001$ , two-tailed.

The reliability of each self-report measure used in this study was examined using Cronbach's alpha coefficient (Cronbach, 1951). Results are presented in Table 4, and all the coefficients exceed the conventional .70 benchmark (Bernardi, 1994; Nunnally, 1978). The measures appear to be assessing relatively unidimensional constructs.

Table 4

*Cronbach's Alpha Coefficients for Self-Report*

*Measures*

---

Measure	Alpha	Number of Items
SF-36 Psychological Health	.82	5
Alcohol Use	.75	20
Job Dissatisfaction	.82	12
Perceived Social Support	.97	38
SF-36 Physical health	.87	10
Disease Prevention	.78	3

---

*Descriptive Statistics and Diagnostics*

Table 5 presents descriptive data for all performance and primary self-report variables obtained at baseline and during the follow-up period. Data about depression, mortality, and age at baseline JHPS assessment obtained from other sources (e.g., medical chart review, note from a treating physician, death certificates) in order to construct primary self-report variables were also included.

Table 5

*Descriptive Statistics for Study Variables*

Rorschach Variable <sup>a</sup>	<i>n</i>	Mean	Standard Deviation	Skew	Kurtosis	Minimum/ Maximum
<i>EII-2</i>	412	-.37	1.03	1.02	2.35	-2.74/4.95
<i>Sum FQ-</i>	412	5.62	3.68	1.00	.99	0/22
Primitive Contents	412	7.81	4.51	1.17	2.30	0/31
<i>GHR</i>	412	4.35	2.53	.56	-.07	0/13
<i>PHR</i>	412	4.95	3.68	1.18	2.22	0/24
<i>M-</i>	412	.72	1.22	3.65	26.00	0/13
<i>WSum6</i>	412	6.49	7.29	1.85	4.65	0/45
<i>R</i>	412	34.44	9.87	.03	-1.24	15/50
Criterion Variable <sup>b</sup>	<i>n</i>	Mean	Standard Deviation	Skew	Kurtosis	Minimum/ Maximum
SF-36 Psychological Health Scale <sup>c</sup>	319	.05	.83	-1.61	3.78	-4.14/.76
Alcohol Use Scale <sup>d</sup>	360	8.45	7.61	1.40	3.26	0/53
Job Dissatisfaction Scale <sup>c</sup>	330	-.09	.70	1.28	1.58	-.89/2.70
SF-36 Physical Health Scale <sup>c</sup>	318	-.17	.96	-1.91	3.94	-4.34/.76
Preventive Health Survey	271	4.57	1.75	-.63	-.13	0/7
Social Support	285	4.51	.71	-2.03	4.37	1.21/5
Total Impairment <sup>e</sup>	217	-.27	3.14	.92	.87	-6.16/11.35
Construction Variable <sup>f</sup>	<i>n</i>	Mean	Standard Deviation	Skew	Kurtosis	Minimum/ Maximum
Age at Rorschach Administration	412	24.56	2.21	.61	.03	20/30
Age at First Depression <sup>g</sup>	49	43.16	9.85	-.48	-.86	23/57
Time from Assessment to First Depression <sup>g</sup>	49	18.96	9.99	-.58	-.45	-5/36
Age at Death	173	67.85	13.85	-.97	.56	22/93

Time from Assessment to Death	173	43.06	13.79	-1.01	.421	1/63
Age at First Marriage <sup>h</sup>	61	26.38	4.74	1.80	5.02	19/47
# Years of First Marriage <sup>h</sup>	61	17.41	8.45	-.21	-.47	1/38

*Note.* *n* refers to the number of cases who provided data on that variable.

<sup>a</sup> Administered at baseline. <sup>b</sup> Administered during follow-up period. <sup>c</sup> Computed in Z-score metric.

<sup>d</sup> Mean number of drinks per week. <sup>e</sup> Aggregate variable. <sup>f</sup> Data used to construct variables for the survival analyses.

<sup>g</sup> Before age 58. <sup>h</sup> If divorced before 1988.

Univariate skewness and/or kurtosis for the self-report variables *M-*, *WSum6*, *SF-36 Psychological Health*, and *Alcohol Use* were elevated or depressed; and kurtosis was dramatically elevated for *M-*. However, these variables were not transformed because the impact of non-normality in a distribution decreases as sample size increases, especially as sample sizes surpass 100 or 200 (Tabachnick & Fidell, 2001).

The impact of possible outliers on the Cox regression analyses was explored with the *DFBETA* statistic available in SPSS. For each covariate, *DFBETA* computes how many standard errors each case would increase or decrease the final regression coefficient if it remained in the analysis. A case with a higher absolute *DFBETA* value is likely to be overly influential (Norusis, 2008). Here, the highest absolute value for any case was 0.03, which represents a negligible influence.

Survival analyses assume that the rate of “failures” over time will be the same for any two cases, or any two groups. Psychological or environmental factors might cause one individual to “survive” longer than another; but it is assumed that, once failures begin for the first person, the rate of failing will be the same for the next (Tabachnick & Fidell, 2001). To determine if this proportional hazards assumption was met here, partial

residual plots (P-plots) were generated and examined. P-plots graphically represent the relationship between survival time and the difference values between observed and expected scores on the covariate for each uncensored case. The proportional hazards assumption is upheld if cases are distributed randomly about the horizontal reference line. Figures 2, 3, and 4 display p-plots for EII-2 partial residuals against survival time for the depression, mortality, and divorce analyses, respectively. The p-plots suggest that the proportional hazards assumption was grossly met in each analysis.

Figure 2

*Scatterplot of EII-2 Partial Residuals by Follow-Up Years in the Depression Analysis*

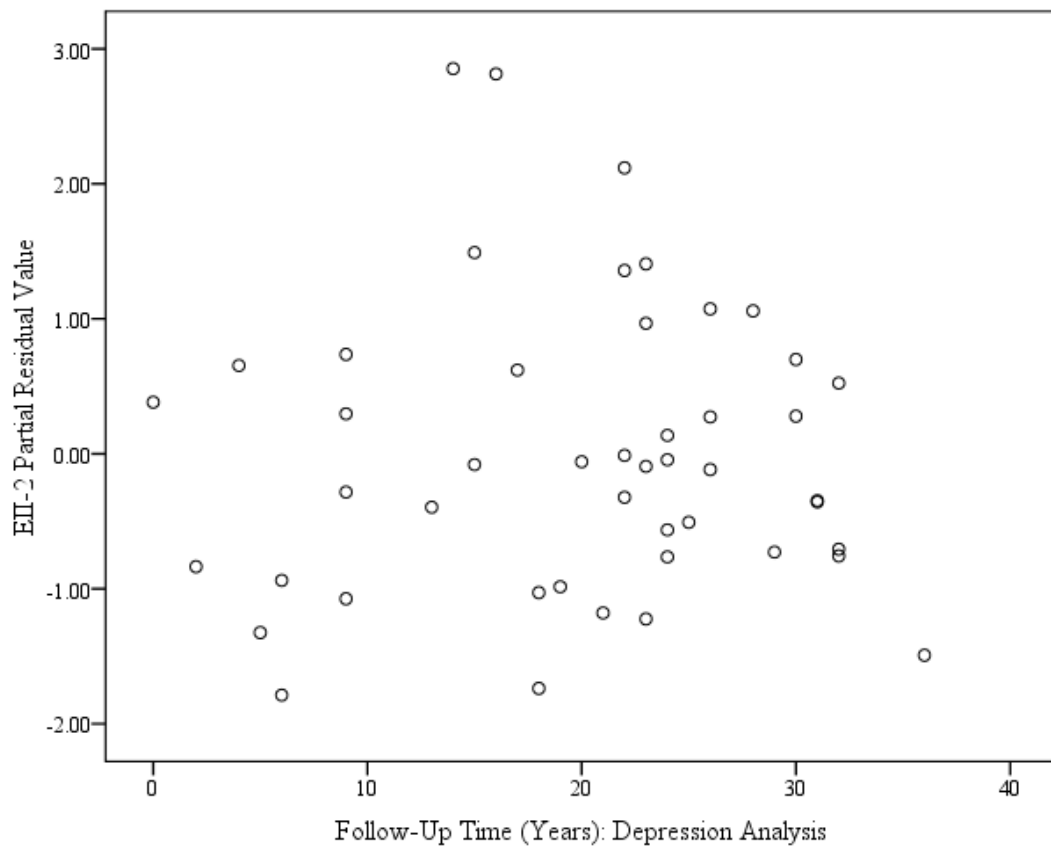


Figure 3

*Scatterplot of EII-2 Partial Residuals by Follow-Up Years in the Mortality Analysis*

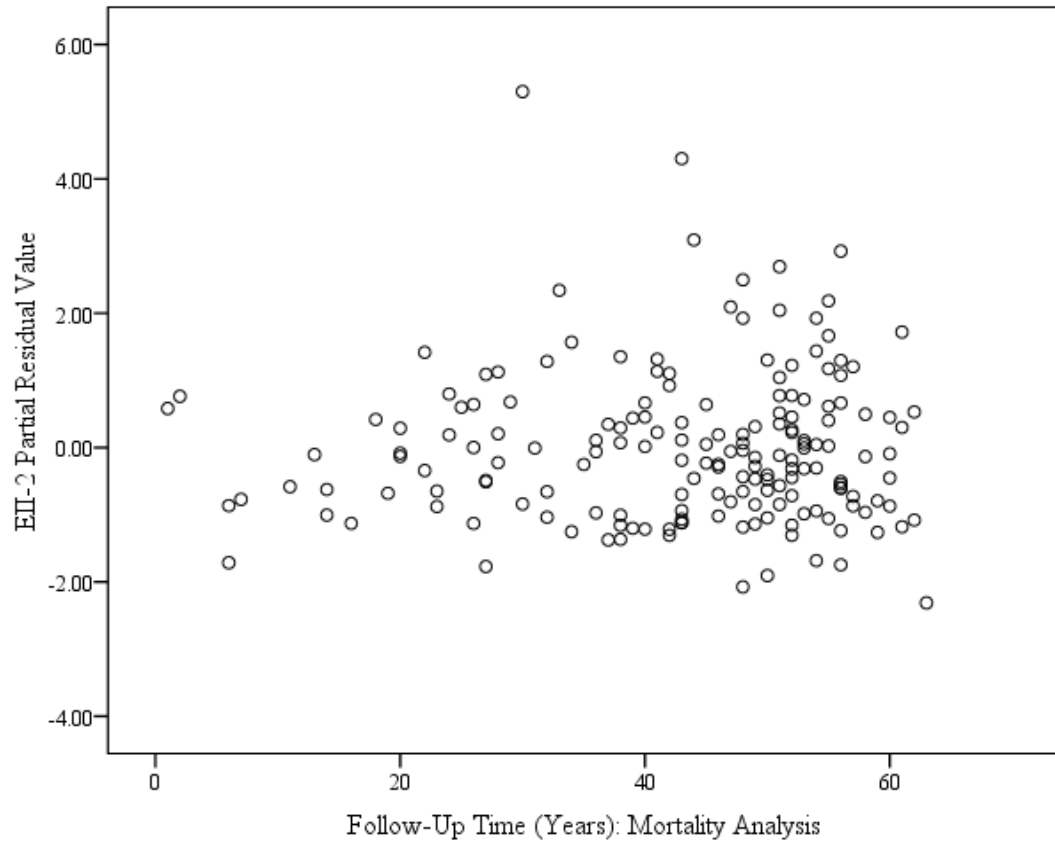
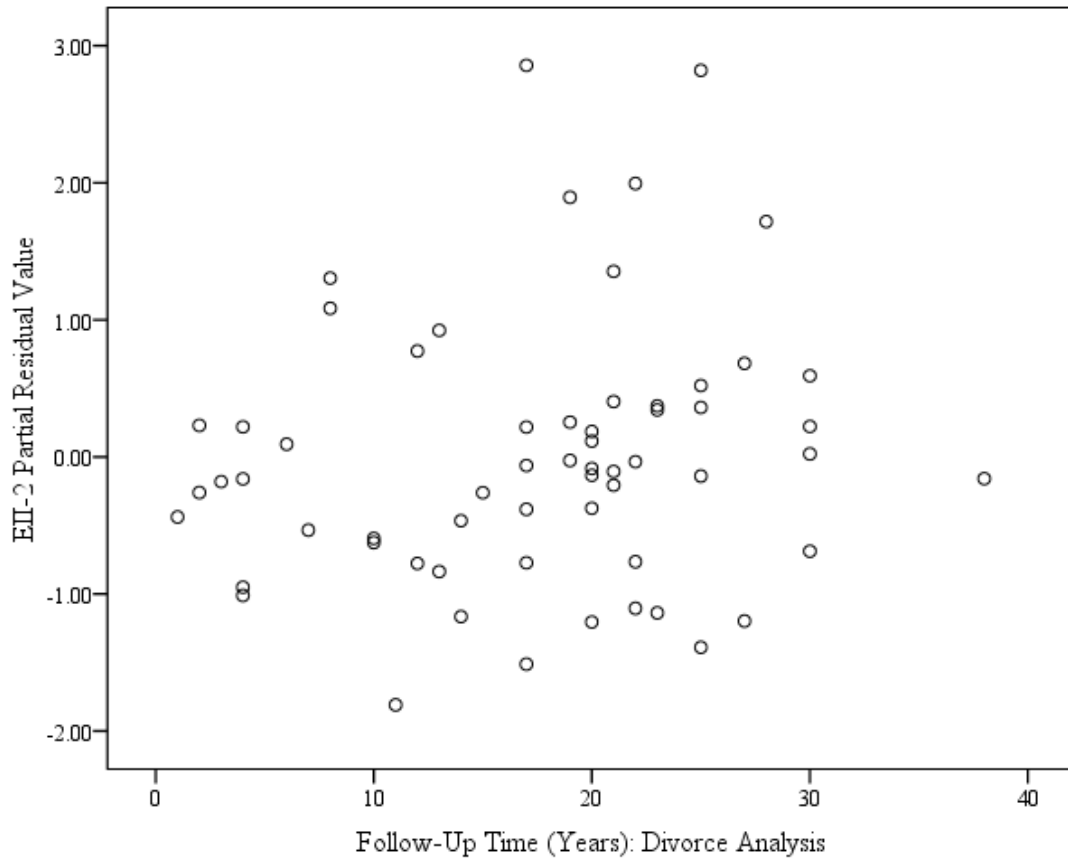




Figure 4

*Scatterplot of EII-2 Partial Residuals by Follow-Up Years in the Divorce Analysis*



The possible existence of outliers was also explored in the linear and multivariate linear hierarchical regressions. The Leverage statistic indicates how far a specific case on a predictor diverges from the mean of the remaining cases. Leverage values above  $3(p + 1)/N$  warrant more attention, i.e., further diagnostics (Howell, 2002). In the multivariate linear regression examining the ability of EII-2 components to predict psychological health, 12 cases had Leverage values above .06 [ $(3[5 + 1])/319 = .06$ ]. If another outlier statistic identifies these cases as well, they would be considered overly influential. Cook's Distance, like DFBETA, estimates the influence of a single case on the regression

model, and values approaching 1.00 are considered high (Howell, 2002). None of the cases in this analysis exceeded .093. This suggests that the cases with high Leverage values are not particularly influential.

The same outlier exploration was conducted for the linear regression examining if overall EII-2 scores can predict psychological health. This time, 11 cases had Leverage values exceeding the conventional cutoff, but Cook's Distance values remained below .08. No outliers were overly influential.

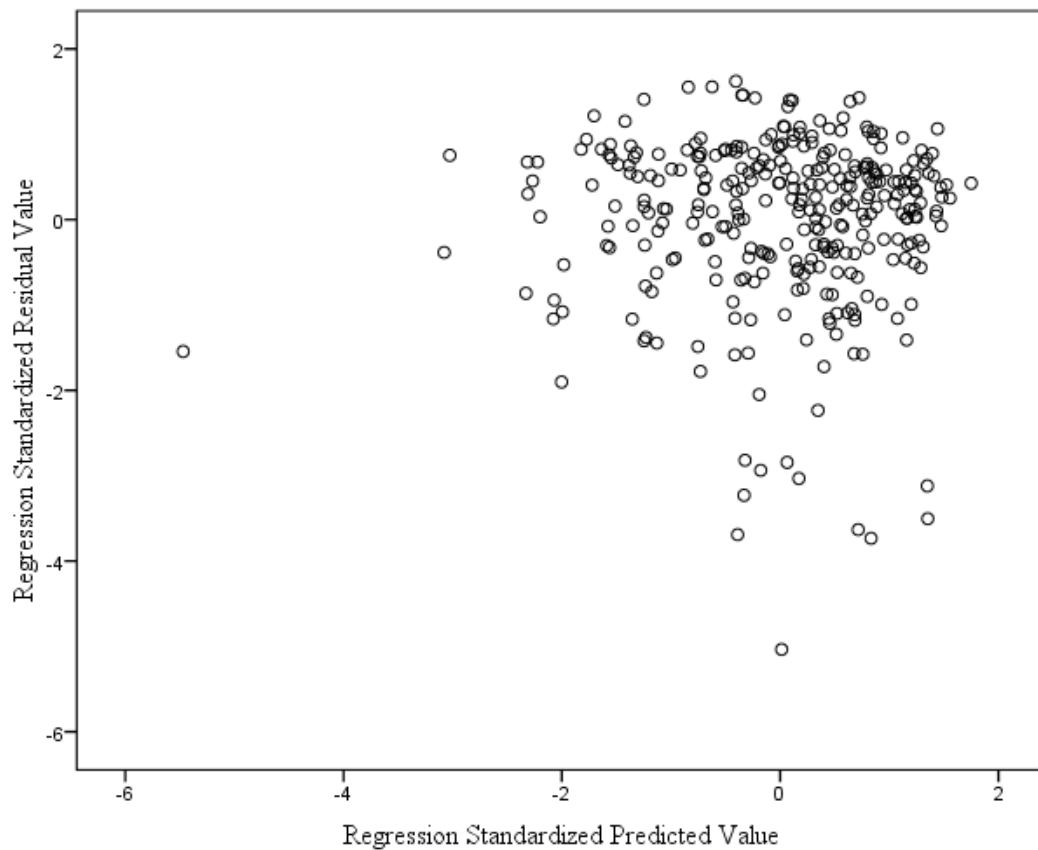
Multicollinearity refers to the degree to which two predictors have a linear relationship. Stronger relationships among predictors make one of the predictors increasingly redundant because it ceases to account for unique variance in the DV. The Tolerance statistic estimates how much variance in a predictor is unique, and not accounted for by other predictors. Values can range from 0.0 to 1.0; lower values indicate that a predictor has little uniqueness. Tolerance values for the multivariate linear regression examining the ability of EII-2 components to predict psychological health ranged from .58 to .96; this suggests multicollinearity was not an issue (Tabachnick & Fidell, 2001).

Another assumption of regression is that the variability of error variance in the IVs is the same at each level of the DV. If the assumption is met (homoscedasticity), the regression equation will be predicting with equal accuracy across levels of the DV. This assumption can be tested by plotting standardized residual values of the variables against regression standardized predicted values. Figure 5 below presents a scatterplot of standardized residual values for the EII-2 components against the regression standardized predicted scores for the SF-36 Psychological Health scale. The resulting distribution

indicates heteroscedasticity among the variables; however, as with multicollinearity, heteroscedasticity is attenuated by the large sample size.

Figure 5

*Scatterplot of Standardized Residual Values of EII-2 Components by the Standardized Predicted Scores for the SF-36 Psychological Health Scale*



*Hypothesis 1: Predictive Validity (Depression)*

Table 6 provides results from the Cox regression analysis examining the ability of the overall EII-2 score to predict time to first depressive episode up to age 57. Two cases reported having had depressive episodes before the JHPS baseline assessment, so their

data were dropped. The final  $n$  for this analysis was 410; 47 cases experienced a first depressive episode before age 57. Data from this analysis indicate that odds ratio for the EII-2 was 1.07, so that for each one point increase in EII-2 score (increased ego impairment), the odds of having a depressive episode before age 57 increased by 7%. However, the overall regression model was not significant, Chi-square (1,  $n = 410$ ) = .26,  $p = .61$ . The EII-2 did not demonstrate a significant effect on time to first depressive episode,  $R = .03$ . Hypothesis 1 was not supported.

Table 6

*Cox Regression Results for EII-2 as Predictor of Time to First Depressive Episode*

Variable	<i>B</i>	<i>SE</i>	<i>df</i>	<i>R</i>	<i>R</i> <sup>2</sup>	<i>Sig.</i>	<i>Exp(B)</i>	<i>95% CI for Exp(B)</i>
<i>EII-2</i>	.07	.14	1	.03	.00	.61	1.07	.88 – 1.30

*Note.*  $n = 410$ ; CI refers to Confidence Interval.

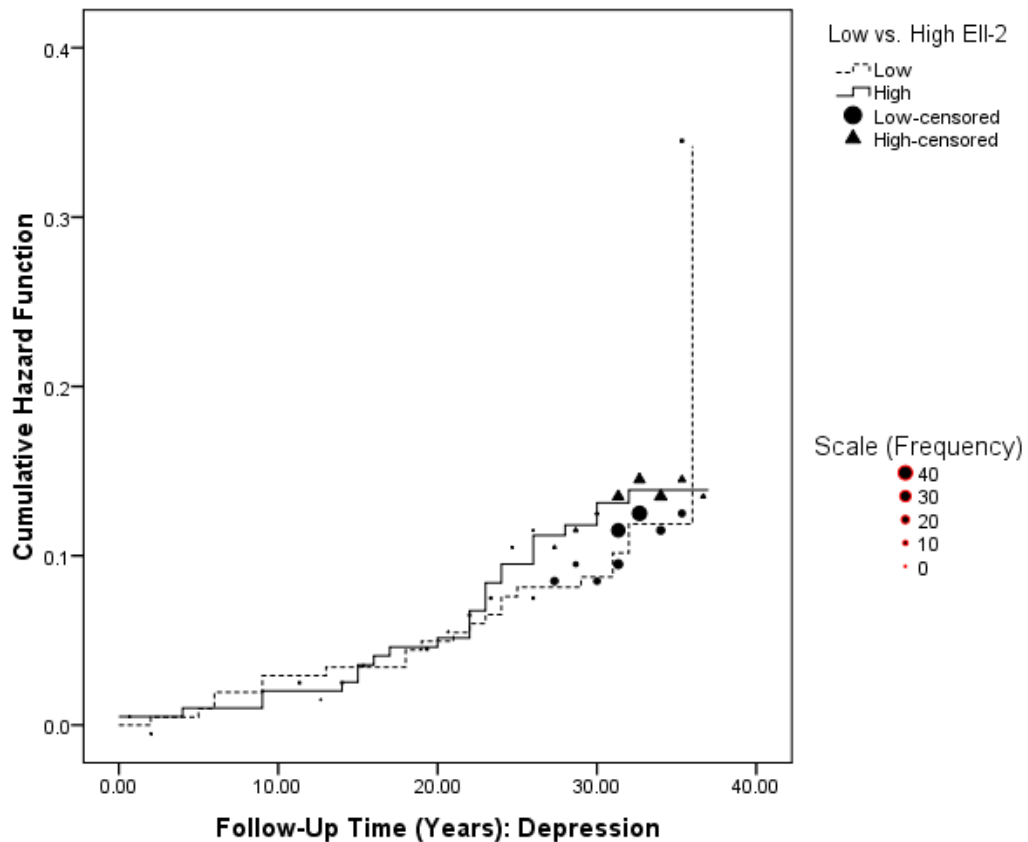
*Hypothesis 2: Predictive Validity (Depression)*

Figure 6 provides results from the Kaplan-Meier analysis. The mean “survival” time from age at JHPS baseline assessment to first depressive episode was 34.84 years for high scorers ( $n = 202$ ), and 34.22 years for low scorers ( $n = 208$ ). The number of high scorers and low scorers who experienced their first depressive episode was 25 and 22, respectively. It was hypothesized that individuals who scored high on the EII-2 would exhibit significantly greater cumulative incidence of depression than low scorers over

time. This was not supported; the hazard functions for high and low EII-2 scorers were not significantly different, Chi-square (1,  $n = 410$ ) = .20,  $p = .66$ ,  $Phi = .02$ .

Figure 6

*Cumulative Hazard Function Plot of Depression by Follow-Up Time in Years for High and Low EII-2 Scorers*



*Hypothesis 3: Predictive Validity (Psychological Health)*

It was hypothesized the overall EII-2 score would significantly predict SF-36 Psychological Health scale scores. The SF-36 questionnaires were administered in 1992, 1995, 1998, 2001, and 2005. The data, presented in Table 7, indicate that the total EII-2 score did not predict psychological health better than chance factors,  $R^2 = .00$ ,  $F(1, 317) =$

1.43,  $p = .23$ . The effect size was very small,  $R = .07$ , but the correlation between the EII-2 and psychological health was negative, which indicates that the association in the regression was in the expected direction. Still, almost no linear relationship appears to exist between the EII-2 and the SF-36 Psychological Health scale. Hypothesis 3 was not supported.

Table 7

*Hierarchical Linear Regression Results for Total EII-2 Score as Predictor of SF-36 Psychological Health Scale Score*

Predictor	<i>B</i>	<i>SE</i>	<i>R</i>	$R^2$	<i>df</i>	<i>F</i>	<i>Sig.</i>
EII-2 Total Score	-.06	.22	.07	.00	1, 317	1.43	.23

*Note.*  $n = 319$ .

*Hypothesis 4: Predictive Validity (Psychological Health)*

It was hypothesized that the interpersonal components of the EII-2 (*GHR*, *PHR*) would predict the SF-36 Psychological Health scale score better than the thought disorder component (*WSum6*), which in turn, would predict better than a combined interpersonal / perceptual accuracy component (*M-*); and all of these components would predict better than Primitive Contents. As noted, the SF-36 Psychological Health scale was administered in 1992, 1995, 1998, 2001, and 2005. Table 7 provides results from the multiple linear regression analysis. In order to lessen the impact of the number of responses per protocol, each EII-2 component was divided by the number of responses

before the analysis. Only 319 cases reported data for the SF-36 Psychological Health scale over the years. The results are consistent with the expectation that Primitive Contents would not account for a significant amount of the variance in psychological health,  $R = .10$ ,  $R^2 = .01$ ,  $F(1, 317) = 3.09$ ,  $p = .08$ . However, contrary to expectations, none of the other EII-2 components fared better. The GoodHR/PoorHR combination did not account for a significant portion of variance after controlling for other EII-2 components,  $R = .12$ ,  $R^2 \text{ change} = .00$ ,  $F(2, 313) = .74$ ,  $p = .48$ ; *WSum6* was not a significant predictor of psychological health after controlling for *M-* and Primitive Contents,  $R = .10$ ,  $R^2 \text{ change} = .00$ ,  $F(1, 315) = .02$ ,  $p = .88$ ; *M-* did not improve upon Primitive Contents as a predictor,  $R = .10$ ,  $R^2 \text{ change} = .00$ ,  $F(1, 316) = .01$ ,  $p = .92$ . Hypothesis 4 was not supported.

Table 8

*Hierarchical Linear Regression Results for EII-2 Components as Predictors of SF-36 Psychological Health Scale Scores*

Step	Variable	B	SE	R	R <sup>2</sup>	Change Statistics			
						R <sup>2</sup> Change	F Change	df	Change Sig.
1	Primitive Contents	-.65	.37	.10	.01	.01	3.09	1, 317	.08
2	<i>M-</i>	-.16	1.65	.10	.01	.00	.01	1, 316	.92
3	<i>WSum6</i>	-.03	.22	.10	.01	.00	.02	1, 315	.88
4	<i>GoodHR</i>	-.21	.64	.12	.01	.00	.74	2, 313	.48
	<i>PoorHR</i>	-.73	.63						

Note.  $n = 319$ . EII-2 components divided by R before analysis.

*Hypothesis 5: Predictive Validity (Mortality)*

Table 9 provides results from the Cox regression analysis examining the ability of the overall EII-2 score to predict time to death. The final sample size for this analysis was 412; 177 of these individuals died during the follow-up period. Data from this analysis indicated that the overall regression model was not significant, Chi-square (1,  $n = 412$ ) = .01,  $p = .93$ . The EII-2 did not demonstrate a significant effect on time to death,  $R = .00$ . The odds ratio for the EII-2 was 1.01, which suggests that, for each one point increase in EII-2 score (increased ego impairment), the odds of dying increased by 1%. Hypothesis 5 was not supported.

Table 9

*Cox Regression Results for EII-2 as Predictor of Time to Death*

Variable	<i>B</i>	<i>SE</i>	<i>df</i>	<i>R</i>	<i>R</i> <sup>2</sup>	<i>Sig.</i>	<i>Exp(B)</i>	<i>95% CI for Exp(B)</i>
<i>EII-2</i>	.01	.07	1	.00	.00	.93	1.01	.90 – 1.12

*Note.*  $n = 412$ ; CI refers to Confidence Interval.

*Hypothesis 6: Predictive Validity (Mortality)*

Figure 7 presents results from the Kaplan-Meier analysis. On average, high EII-2 scorers ( $n = 204$ ) lived 54.12 years after completing the JHPS baseline assessment, while low scorers ( $n = 208$ ) lived 54.05 years. Essentially the same number of high scorers ( $n = 89$ ) died during the follow-up period than did low scorers ( $n = 88$ ). It was hypothesized that individuals who scored high on the EII-2 would exhibit significantly greater

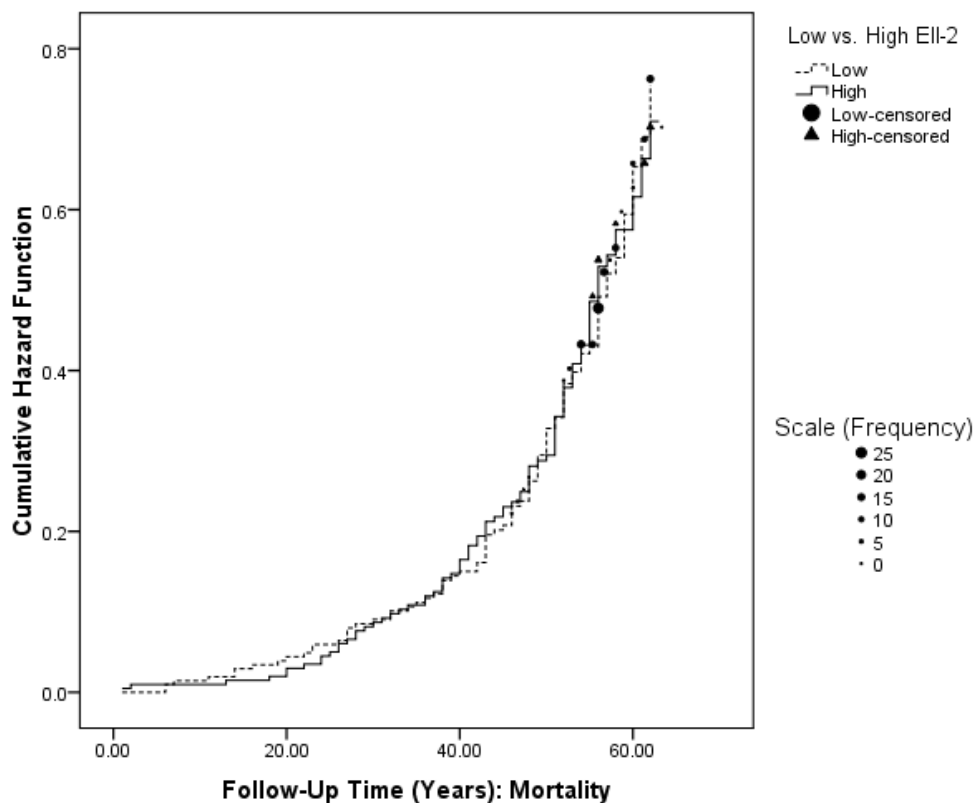


cumulative incidence of mortality than low scorers over time. This was not supported; the hazard functions for high and low EII-2 scorers were not significantly different, Chi-square (1,  $n = 412$ ) = .00,  $p = .98$ ,  $\phi = .00$ .

Figure 7

*Cumulative Hazard Function Plot of Mortality by Follow-Up*

*Time in Years for High and Low EII-2 Scorers*



*Hypothesis 7: Predictive Validity (Divorce)*

It was hypothesized that EII-2 scores would predict the cumulative incidence of divorce over time. Results from the Cox regression analysis are in Table 10. Of the 269 total individuals who reported information about marriage status during the follow-up

period, 61 experienced their first divorce by 1987. Here, the overall regression model was not significant, Chi-square (1,  $n = 269$ ) = 1.12,  $p = .30$ , and the EII-2 did not show a strong association with time to first divorce,  $R = .05$ . The odds ratio was 1.13, so the odds of divorcing increased by 13% for every one point increase in the EII-2 score (increased ego impairment). However, hypothesis 7 was not supported.

Table 10

*Cox Regression Results for EII-2 as Predictor of Time to First Divorce*

Variable	<i>B</i>	<i>SE</i>	<i>df</i>	<i>R</i>	<i>R</i> <sup>2</sup>	<i>Sig.</i>	<i>Exp(B)</i>	<i>95% CI for Exp(B)</i>
<i>EII-2</i>	.12	.12	1	.05	.00	.29	1.13	.96 – 1.34

*Note.*  $n = 269$ ; CI refers to Confidence Interval.

*Hypotheses 8 – 13: Construct Validity*

The results for all bivariate correlations in the concurrent validity analyses are presented in Table 11. All self-report measures in these analyses were administered during the follow-up period. It was hypothesized that higher EII-2 scores (more impairment) would correlate significantly and positively with alcohol use, job dissatisfaction, and the aggregate variable based on all variables but mortality, but significantly and negatively with perceived social support from others, physical health, preventive health practices. The correlation between the EII-2 and the Total Impairment variable was significant ( $r = .16$ ,  $n = 217$ ,  $p = .02$ ), so hypothesis 13 was supported; no other construct validity hypotheses were supported. The correlation between the EII-2 and the Alcohol Use

variable was in the hypothesized direction, but the magnitude was negligible ( $r = .05$ ,  $n = 360$ ,  $p = .31$ ). The Prevention ( $r = -.07$ ,  $n = 271$ ,  $p = .23$ ), Job Dissatisfaction ( $r = .05$ ,  $n = 330$ ,  $p = .32$ ), and Physical Health ( $r = -.02$ ,  $n = 318$ ,  $p = .65$ ) variables also related to the EII-2 in the direction hypothesized; but their magnitudes were not significant. Perceived social support ( $r = .02$ ,  $n = 285$ ,  $p = .79$ ) correlated with the EII-2 in the opposite direction than expected. These data suggest that EII-2 scores have no meaningful association with the concurrent validity variables when they are considered separately; but the relationship strengthens considerably when these variables are aggregated into a single variable.

Table 11

*Bivariate Correlations between the EII-2 and Self-Report and Aggregate Construct Validity Variables*

Variable	<i>r</i>	<i>p</i>	<i>n</i>
Alcohol Use: Mean # Drinks Per Week	.05	.31	360
Job Dissatisfaction	.05	.32	330
Perceived Social Support	.02	.79	285
Physical Health	-.02	.65	318
Preventive Health Practices	-.07	.23	271
Total Impairment	.16	.02	217

*Note.* Significance test (*p*) was two-tailed.

A correlation matrix of the primary study variables is presented below (Table 12). Excluding correlations between the Total Impairment variable and its individual components, 13 of the 47 (28%) associations were significant at the  $p < .05$  level. As expected, ego impairment (EII-2) was associated with Total Impairment; Total

Impairment was associated with death; psychological health was associated with physical health, less job dissatisfaction, depression, and death; depression and death were associated; job dissatisfaction was associated with less perceived social support; perceived social support was associated with less preventive health activity; and poorer physical health was associated with depression and death.

Table 12

*Intercorrelations between Primary Study Variables*

Variable	EII-2	SF-Psy	AU	JD	PSS	SF-Health	PHP	Depr	Death	Divorce	Total Imp.
EII-2	-	-.07	.05	.05	.02	-.02	-.07	.03	.02	.08	.16*
Psychological											
Health (SF-Psy) <sup>a</sup>	<i>n</i> = 319	-	-.09	-.28**	.11	.39**	.06	-.19**	-.15**	.05	-.57**
Alcohol Use (AU) <sup>b</sup>	<i>n</i> = 360	<i>n</i> = 316	-	.06	.04	-.05	-.03	.01	.05	.11	.40**
Job Dissatisfaction (JD) <sup>c</sup>	<i>n</i> = 330	<i>n</i> = 311	<i>n</i> = 329	-	-.20**	-.12*	-.09	.18*	-.05	-.00	.47**
Perceived Social											
Support (PSS) <sup>d</sup>	<i>n</i> = 285	<i>n</i> = 285	<i>n</i> = 284	<i>n</i> = 279	-	.07	.15*	-.01	-.03	-.09	-.47**
Physical Health											
(SF-Health) <sup>a</sup>	<i>n</i> = 318	<i>n</i> = 318	<i>n</i> = 315	<i>n</i> = 310	<i>n</i> = 285	-	.07	-.12*	-.32**	-.08	-.51**
Preventive Health											
Practices (PHP) <sup>e</sup>	<i>n</i> = 271	<i>n</i> = 271	<i>n</i> = 271	<i>n</i> = 270	<i>n</i> = 258	<i>n</i> = 271	-	-.04	-.06	-.04	-.44**
Depression (Depr):											
Dichotomous <sup>f</sup>	<i>n</i> = 412	<i>n</i> = 319	<i>n</i> = 360	<i>n</i> = 330	<i>n</i> = 285	<i>n</i> = 318	<i>n</i> = 271	-	.12*	-.01	.42**
Death: Dichotomous <sup>g</sup>	<i>n</i> = 412	<i>n</i> = 319	<i>n</i> = 360	<i>n</i> = 330	<i>n</i> = 285	<i>n</i> = 318	<i>n</i> = 271	<i>n</i> = 412	-	-.04	.18**
Divorce: Dichotomous <sup>h</sup>	<i>n</i> = 269	<i>n</i> = 253	<i>n</i> = 269	<i>n</i> = 269	<i>n</i> = 231	<i>n</i> = 252	<i>n</i> = 225	<i>n</i> = 269	<i>n</i> = 269	-	.38**
Total Impairment <sup>i</sup>	<i>n</i> = 217	<i>n</i> = 217	<i>n</i> = 217	<i>n</i> = 217	<i>n</i> = 217	<i>n</i> = 217	<i>n</i> = 217	<i>n</i> = 217	<i>n</i> = 217	<i>n</i> = 217	-

<sup>a</sup> Administered in 1992, 1995, 1998, 2001, and 2005. <sup>b</sup> Administered in 1978, 1986, 1989, 1993, 1997, and 2003. <sup>c</sup> Administered in 1987, 1991, and 1996.

<sup>d</sup> Administered in 1997. <sup>e</sup> Administered in 1997. <sup>f</sup> Computed as Depression before age 58: no = 0, yes = 1. <sup>g</sup> Computed as Dead: no = 0, yes = 1. <sup>h</sup> Computed as Divorce before 1988: no = 0, yes = 1. <sup>i</sup> Aggregate variable.

\*  $p < .05$ , two-tailed. \*\*  $p < .01$ , two-tailed.

### *Secondary Analyses*

As noted, the possible influence of examiner differences and protocol complexity were controlled in secondary analyses for all of the criterion variables. Entering these covariates in regression analyses did not substantively alter the findings for any variable.

## Chapter 4

### Discussion

The primary purpose of this study was to examine the predictive validity of the EII-2, a measure of ego impairment generated from Exner's (2003) CS for the Rorschach Inkblot Test. Consistent with Viglione et al.'s (2003) suggestion to use behavioral markers as criterion variables to assess EII-2 constructs like information processing, decision making, and adaptive functioning, survival analyses were conducted to determine if baseline EII-2 scores could predict the cumulative incidence of depression and mortality in a long-term, prospective cohort (the assessment of depression included self-report). Predictive validity was also examined using follow-up self-report data about divorce and psychological health. A secondary purpose of this study was to follow recommendations by previous authors (Perry & Viglione, 1991; Perry et al., 1992; Viglione, Perry, Jansak, et al., 2003) to examine the relative contributions of the individual EII-2 components in predicting meaningful criterion variables, e.g., psychological health. Finally, construct validity was examined by relating the EII-2 overall score to follow-up questionnaire data regarding alcohol use, job dissatisfaction, physical health, the use of preventative health services, and perceived social support, and an aggregate variable constructed from these variables.

Most of the hypotheses in this study were not supported, which weakens the overall case for the predictive validity of the EII-2 as a measure of ego impairment. The



effect size from the Cox regression examining the EII-2 as predictor of the cumulative incidence of depression was negligible, though in the expected direction. This suggests that the EII-2 scores had no relationship with the rate at which a first depressive episode occurred during the follow-up period. In the other survival analyses, the EII-2 was unable to predict time to death and time to first divorce, but again, the relationships were in the expected direction. Similarly, the EII-2 and its components did not predict psychological health in the linear regressions, but relationships were in the expected directions (when not zero). None of the correlations with the separate self-report variables reached significance; but only the correlation with social support was in the unexpected direction.

In contrast, the results supported hypothesis 13, which predicted that the correlation between the EII-2 and the Total Impairment aggregate variable would be positive and significant. This supports the construct validity of the EII-2 by indicating that the baseline EII-2 was a precursor to a specific group of physical, psychosocial, and psychological outcomes in later life that would be expected if the EII-2 measures an underlying, core personality structure responsible for adaptive functioning in the environment. At first glance, this finding may seem inconsistent with the fact that the relationships between the EII-2 and the individual components of the aggregate variable (depression, divorce, alcohol use, job dissatisfaction, social support, physical health, psychological health, preventive health) were of negligible magnitude. However, aggregation involves converting scores to the same standardized metric, then “accumulating” them into a single score. In this process, measurement error associated with each score is averaged out, so that the final score is a more stable estimator of the target phenomenon than any of the individual ones (Rushton, Brainerd, & Pressley, 1983).

In turn, this increases the power to detect a significant effect or relationship if it truly exists (Lipsey & Wilson, 2001).

The relationship between the EII-2 and the Total Impairment aggregate variable was consistent with other studies that support the construct validity of the EII-2. Adrian and Kaser-Boyd (1995), for example, found that EII scores correlated significantly with MMPI Scale 8 ( $r = .25, n = 85, p < .05$ ), and several of its subscales, in a diverse group of psychiatric patients. Perry and Braff (1994) found that higher EII – *HEV* scores were associated with more problems on information-processing tasks in a sample of individuals with schizophrenia (Spearman's  $r = -.42, n = 35, p < .01$ ). Further, the EII-2's relationship with the aggregate variable found here is consistent with a recent meta-analysis by Roberts, Kuncel, Shiner, Caspi, and Goldberg (2007). In that study, Big Five personality traits were associated with mortality, divorce, and occupational attainment in longitudinal studies. Follow-up periods ranged from 1 year to 78 years. Conscientiousness ( $N = 4$  studies ;  $r = -.09, CI = -.12$  to  $-.05$ ) and Extraversion ( $N = 6; r = -.07, CI = -.11$  to  $-.03$ ) were associated with less mortality. Negative Emotionality ( $N = 12; r = .05, CI = .02$  to  $.08$ ) and Hostility/Disagreeableness ( $N = 19; r = .04, CI = .02$  to  $.06$ ) were associated with shorter life spans. These effect sizes were all in the small range for mortality. Across 13 studies, Neuroticism and similar traits were associated with divorce ( $r = .17, CI = .12$  to  $.22$ ), while Agreeableness ( $r = -.18, CI = -.27$  to  $-.09$ ) and Conscientiousness ( $r = -.13, CI = -.17$  to  $-.09$ ) related to less divorce.

Regarding the analyses with individual variables, the EII-2 results found here were not consistent with previous EII/EII-2 research. In Perry and Viglione's (1991) original EII-2 article, for example, the EII-2 was a highly significant predictor of melancholic

depressed individuals' responses to tricyclic antidepressants across the 9 week treatment period. Further, Perry, McDougall, and Viglione (1995) found that the EII-2 demonstrated impressive temporal stability over a 5 year period (Spearman's  $r = .68$ ,  $p < .01$ ), which suggests it is sensitive to a stable core structure over long periods of time. However, it may be that predicting the behavioral manifestations of ego functioning across a few years is quite different than predicting across several decades. This might be the case, for example, if the EII-2 is only sensitive to state-level ego functioning, and therefore not capable of capturing the more trait-level, core ego structure across many years, as it purports to.

The prediction analyses also appear to be inconsistent with some research focusing on personality characteristics as predictors of long-term behavioral and health outcomes. Kawachi et al. (1996) found that responses to the MMPI-2 Anger Content Scale predicted coronary heart disease (CHD) during a 7 year follow-up period. Individuals who scored 0 or 1 on the scale had increased risk of CHD (Relative Risk [RR] = 2.76, 95% CI = .082 to 9.26) compared to non-angry respondents, while individuals who scored between 5 and 14 more than doubled their risk (3.58, CI = 1.08 to 11.9). In the JHPS literature, Graves et al. (1994) found that individuals who were identified as anxious and inhibited at baseline were at higher risk of early mortality (RR= 1.56, 95% CI = 1.00 to 2.44) when compared to the emotionally stable group.

Another possible reason for the inconsistency with Perry and Viglione's (1991) data, and the mild associations between the EII-2 and the JHPS variables, is that the EII-2 may be an imprecise measure of ego functioning at the healthier end of the health-pathology spectrum (Viglione et al., 2003). The mean EII-2 score for the JHPS cohort was

“optimal” (healthiest) according to Viglione et al.’s (2003) suggested interpretive ranges for EII-2 scores. The mean EII-2 score for non-depressed cohort members was not significantly different from the mean score of members who would later become depressed,  $t(410) = .68, p = .50, r = .03$ . Further, means for both of these groups were in the “no impairment” range for the EII-2 (-.30 - +.20) according to Viglione et al. (2003). The EII-2 might not capture the true variability in ego functioning in this cohort, and any true relationships between EII-2 scores and outcome variables might not be detected.

The secondary analyses indicated that controlling for examiner influences made no difference in the EII-2’s predictive ability. However, there are compelling reasons to continue exploring the examiner difference issue. As noted, the JHPS Rorschach protocols were administered using Beck’s (1944) procedure, then recoded for the present study using Exner’s (2003) CS. Beck’s administration procedure has inquiry instructions that are less specific and detailed than Exner’s. This may have contributed to the examiner differences described. Some examiners appeared to record response and inquiry material verbatim at times. When these examiners’ protocols were recoded for this project, plenty of rich material was available for coding CS cognitive Special Scores if the examinee presented with disturbed thought process. As noted, the sum of cognitive Special Scores comprises the *WSum6* component of the EII-2. Other JHPS examiners, however, recorded little more than location codes in the inquiry column, so thought disturbance, if it was exhibited by the examinee, was not captured during the administration. One of these original coders administered close to 20% of all the individual Rorschachs. This would likely result in substantial underestimations of thought disturbance and ego impairment in this cohort, which in turn, would obscure any true relationships between the EII-2 and the

outcome variables. When the difference in the average number of responses per protocol is controlled for, the JHPS cohort's mean *WSum6* score (Mean/*R* = .18) was lower (healthier) than the international normative reference sample's mean (M/*R* = .34; Meyer, Erdberg, & Shaffer, 2007). Here, one of the JHPS cohort's healthiest-looking components (*WSum6*) relied most heavily on detailed inquiry. Not surprisingly, this translated into significantly healthier total EII-2 scores for cohort members (M = -.37, SD = 1.03) versus reference sample members (M = -.15, SD = .95), although the effect size of this difference was small ( $p < .001$ ;  $r = -.06$ ). This pattern of departures from the international reference sample may suggest that the coding/inquiry issue created an underestimation of the *WSum6* variable, and perhaps final EII-2 scores. These data contrast with results from the secondary analyses which suggest that examiner differences had little influence on the outcomes of this study. However, it may be that examiners differed from each other along a continuum of amount of inquiry material recorded, so that the ordering of examiners – from those who recorded more inquiry material to those who recorded very little – related with *WSum6* and the EII-2 score in a linear way. Such a relationship might be robust, but not emerge when analyses focus on individual examiners, or the group of examiners in random order. A major limitation of this study was that the possible ordering effects of the examiners were not systematically examined in the primary analyses. This would be an interesting future direction for research using the CS-scored JHPS Rorschach protocols.

Some of the variables in this study raised questions about the results. The excellent alpha coefficients reported earlier indicate that the self-report variables used here were internally consistent, and measuring unidimensional constructs. However, these variables tended not to correlate to each other, and to other study variables, in expected ways. Mean

number of alcohol drinks per week, for example, had no relationship with life/death status ( $r = .04, p = .42, n = 347$ ), depression status ( $r = -.01, p = .80, n = 347$ ), physical health concerns ( $r = -.04, p = .54, n = 193$ ), or age at death ( $r = -.08, p = .36, n = 133$ ). Further, the ‘age of first depression’ variable indicated that the mean age of onset was 43.16 for cohort members who became depressed before age 58. However, the mean age for depression onset for all cohort members in this study was 56.32, an unusually high age. It is notably higher than the typical age-range of onset of depression (25 – 44) reported by the Diagnostic and Statistical Manual of Mental Disorders, Text Revision (DSM-IV-TR; American Psychological Association, 2000). These variable relationships and values were unexpected, and they raise suspicions about validity of the variables.

The JHPS’ time-to-event data permitted the testing of hypotheses about the predictive validity of the EII-2 in novel ways, with survival analyses. As noted, the DV in survival analyses is the time it takes someone to reach a specified event for the first time; and after having reached it once, they are taken out of the analysis (Norusis, 2008). The results of the current study suggest that, at least for white, professional males in the upper socioeconomic bracket, the degree of ego impairment in young adulthood as measured by the EII-2 has little or no influence on the amount of time that will pass before a first depressive episode, death, or first divorce occurs – providing they can and do occur. In other words, the EII-2 may not measure what it purports to measure. It may be, however, that the EII-2 could demonstrate more predictive power if cases did not drop out of the analysis after reaching the event for the first time (not when the event is death). Ego impairment, for instance, might influence the number of events during follow-up, or the times between repeated events, more than it influences the time to the first event. There

are Cox regression models designed to analyze repeated occurrences of a target event during a certain period of time (Garson, 2008). Any revised version of this study would do well to utilize a repeated events Cox regression model, provided sufficient data become available.

As noted, some data suggest that this variable (cognitive flexibility, motivation, problem-solving) has demonstrated the ability to moderate the Rorschach's predictive power (Dean et al., 2007; Morgan & Viglione, 1992). However, in the secondary analyses, it had no influence on the DVs. This is notable given that the complexity variable can enhance the Rorschach's predictive power when the number of responses is in the nonoptimal range, for example, above 28 (Dean et al., 2007). The mean number of responses per protocol in the JHPS cohort was 34.44. Still, complexity demonstrated a small and consistent relationship with the EII-2. For example, the correlation between complexity and the EII-2 was .23 ( $n = 412$ ). If the prediction-enhancement power of complexity is examined again with the CS-scored JHPS protocols, it might be fruitful to only include protocols with  $R$  in the nonoptimal range.

One of the most significant limitations of this study is simultaneously one of its greatest strengths. The cohort sample consisted entirely of highly educated, successful, white males (Shaffer et al., 1982). This severely limits the ability to generalize any of the findings herein to other populations. At the same time, the homogeneity of a sample like this virtually eliminates the influence of many possible threats to validity, such as variability in socioeconomic status, education level, ethnicity, race, gender, etc. (Klag et al., 1993).

The current study was the first to use a prospective, longitudinal cohort design to examine the psychometric properties of the EII-2. Unfortunately, most of the results obtained here stand in contrast to much of the published literature on the EII-2. As noted, for example, the EII-2 was found to be a highly significant predictor of melancholic depressed individuals' responses to tricyclic antidepressants across the 9 week treatment period (Perry & Viglione, 1991). But it has also demonstrated very impressive construct validity. Perry et al. (2003) hypothesized, and found, that the EII could discriminate among groups of non-patients ( $n = 66$ ), college students with elevated scores on a self-report measure of magical thinking ( $n = 24$ ), 1<sup>st</sup> degree relatives of individuals with schizophrenia ( $n = 36$ ), individuals with schizotypal personality disorder ( $n = 36$ ), outpatients with schizophrenia ( $n = 33$ ), and inpatients with schizophrenia ( $n = 56$ ), using an analysis of variance with focused linear contrasts,  $F(1, 245) = 49.58, p < .001, r = .41$ . EII scores increased (pathology increased) significantly across these groups along the schizophrenia spectrum. Perry found the same pattern for each of these five EII components: *FQ*- ( $F[1, 245] = 23.81, p < .001, r = .30$ ), *Wsum6* ( $F[1, 245] = 42.80, p < .001, r = .39$ ), *Critical Contents* ( $F[1, 245] = 7.27, p < .01, r = .17$ ), *GHR* ( $F[1, 245] = 23.20, p < .001, r = .29$ ), and *PHR* ( $F[1, 245] = 10.63, p < .001, r = .20$ ). Although the EII-2 did not end up being a significant predictor of the individual variables in the current study, it has elsewhere proven to be sensitive to perceptual difficulties, cognitive disorganization, illogical thinking, and problems with instinct regulation and defensive functions.



## References

- Adams, H. B., & Cooper, G. D. (1962). Three measures of ego strength and prognosis for psychotherapy. *Journal of Clinical Psychology, 18*, 490-494.
- Adrian, C. & Kaser-Boyd, N. (1995). The Rorschach Ego Impairment Index in heterogeneous psychiatric patients. *Journal of Personality Assessment, 65*, 408-414.
- Alexopolous, G. S., Meyers, B. S., Young, R. C., Campbell, S., Silbersweig, D., & Charlson, M. (1997). 'Vascular Depression' hypothesis. *Archives of General Psychiatry, 54*, 915-922.
- Allen, J. G., Coyne, L., & David, E. (1986). Relations of intelligence to ego functioning in an adult psychiatric population. *Journal of Personality Assessment, 50*, 212-221.
- Allport, G. W. (1943). The ego in contemporary psychology. *Psychological Review, 50*, 451-478.
- Alpher, V. S. (1991). Assessment of ego functioning in multiple personality disorder. *Journal of Personality Assessment, 56*, 373-387.
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders, text revision* (4<sup>th</sup> ed.). Washington, DC: American Psychiatric Association.
- Amin, M. G., Tighiourt, H., Weiner, D. E., Stark, P. C., Griffith, J. L., MacLeod, B., Salem, D. N., et al. (2004). Hematocrit and left ventricular mass: The Framingham heart Study. *Journal of the American College of Cardiology, 43*, 1276-1282.

- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7<sup>th</sup> ed.). Upper Saddle River, NJ: Prentice Hall.
- Andreason, N. (1979). Thought, language, and communication disorders. *Archives of General Psychiatry*, *36*, 1315-1321.
- Andreason, N. (1984). *The Scale for the Assessment of Positive Symptoms (SAPS)*. The University of Iowa, Iowa City.
- Andreason, N., & Olsen, S. (1982). Negative and positive schizophrenia: Definition and validation. *Archives of General Psychiatry*, *39*, 789-793.
- Auslander, L. A., Perry, W., & Jeste, D. V. (2002). Assessing disturbed thinking and cognition using the Ego Impairment Index in older schizophrenic patients: Paranoid vs. nonparanoid distinction. *Schizophrenia Research*, *53*, 199-207.
- Axelrod, P., & Kessel, P. (1972). Residual effects of LSD on ego functioning: An exploratory study with the Rorschach test. *Psychological Reports*, *31*, 547-550.
- Baldwin, R. C. (2005). Is vascular depression a distinct sub-type of depressive disorder? A review of causal evidence. *International Journal of Geriatric Psychiatry*, *20*, 1-11.
- Barron, F. (1953). An ego-strength scale which predicts response to psychotherapy. *Journal of Consulting Psychology*, *17*, 327-333.
- Beck, S. J. (1944). *Rorschach's test: I. Basic processes*. New York, NY: Grune & Stratton.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561-571.

- Becker, W. C. (1956). A genetic approach to the interpretation and evaluation of the process-reactive distinction in schizophrenia. *Journal of Abnormal and Social Psychology, 53*, 229-236.
- Bell, M. D., Billington, R. J., & Becker, B. R. (1985). Scale for the assessment of reality testing: Reliability, validity, and factorial invariance. *Journal of Consulting and Clinical Psychology, 53*, 506-511.
- Bell, M., Billington, R., & Becker, B. (1986). A scale for the assessment of object relations: Reliability, validity, and factorial invariance. *Journal of Clinical Psychology, 42*, 733-741.
- Bell, B., Rose, C., Damon, A. (1972). The Normative Aging Study: An interdisciplinary and longitudinal study of health and aging. *Aging and Human Development, 3*, 4-17.
- Bell, M. D., & Zito, W. (2005). Integrated versus sealed-over recovery in schizophrenia: BORRTI and executive functioning. *Journal of Nervous and mental Disease, 193*, 3-8.
- Bellak, L., Chassan, J. B., Gediman, H. K., & Hurvich, M. (1973). Ego function assessment of analytic psychotherapy combined with drug therapy. *Journal of Nervous and Mental Disease, 157*, 465-469.
- Bellak, L., & Hurvich, M. (1969). A systematic study of ego functions. *Journal of Nervous and mental Disease, 148*, 569-585.
- Beres, D. (1956). Ego deviation and the concept of schizophrenia. *Psychoanalytic Study of the Child, 11*, 164-235.

- Bernardi, R. A. (1994). Validating research results when Cronbach's alpha is below .70: A methodological procedure. *Educational and Psychological Measurement, 54*, 766-775.
- Betz, B. J., & Thomas, C. B. (1979). Individual temperament as a predictor of health and premature disease. *Johns Hopkins Medical Journal, 144*, 81-89.
- Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *Clinical Neuropsychologist, 3*, 129-136.
- Block, J., Block, J., & Keyes, S. (1988). Longitudinally foretelling drug usage in adolescence: Early childhood personality and environmental precursors. *Child Development, 59*, 336-355.
- Breuer, J. & Freud, S. (1957). *Studies on hysteria*. New York: Basic Books (Original work published 1895).
- Buelow, G., McClain, M., & McIntosh, I. (1996). A new measure for an important construct: The attachment and object relations inventory. *Journal of Personality Assessment, 66*, 604-623.
- Cadenhead, K. S., Perry, W., & Braff, D. L. (1996). The relationships of information-processing deficits and clinical symptoms in schizotypal personality disorder. *Biological Psychiatry, 40*, 853-858.
- Calkins, M. W. (1915). The self in scientific psychology. *American Journal of Psychology, 26*, 495-524.
- Carroll, B. J., Feinberg, M., Smouse, P. E., Rawson, S. G., & Greden, J. F. (1981). The Carroll Rating Scale for depression: Development, reliability, and validity. *British Journal of Psychiatry, 138*, 194-200.

- Cartwright, R. D. (1958). Predicting response to client-centered therapy with the Rorschach PR Scale. *Journal of Consulting Psychology, 5*, 11-17.
- Chang, P. P., Ford, D. E., Mead, L. A., Cooper-Patrick, L., & Klag, M. J. (1997). Insomnia in young men and subsequent depression. *American Journal of Epidemiology, 146*, 105-114.
- Chang, P. P., Ford, D. E., Meoni, L. A., Wang, N. Y., & Klag, M. J. (2002). Anger in young men and subsequent premature cardiovascular disease. *Archives of Internal Medicine, 162*, 901-906.
- Christopher, J. C., Bickhard, M. H., & Lambeth, G. S. (2001). Otto Kernberg's object relations theory: A metapsychological critique. *Theory & Psychology, 11*, 687-711.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.
- Conte, H. R., Buckley, P., Picard, S., & Karasu, T. B. (1995). Relationships between psychological mindedness and personality traits and ego functioning: Validity studies. *Comprehensive Psychiatry, 36*, 11-17.
- Cox, D. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, 74*, 187-220.
- Cramer, P., & Block, J. (1998). Preschool antecedents of defense mechanism use in young adults: A longitudinal study. *Journal of Personality and Social Psychology, 74*, 159-169.

- Cromie, W. J. (2001). How to be happy and well rather than sad and sick. *Harvard University Gazette*, Issue June 07, 2001. Retrieved October 11, 2007, from <http://www.hno.harvard.edu/gazette/2001/06.07/01-happywell.html>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.
- Crumpton, E., Cantor, J. M., & Batiste, C. (1960). A factor analytic study of Barrons' Ego Strength Scale. *Journal of Clinical Psychology*, *16*, 283-291.
- Dawes, R. M. (1999). Two methods for studying the incremental validity of the a Rorschach variable. *Psychological Assessment*, *11*, 297-302.
- Dean, K. L., Viglione, D. J., Perry, W., & Meyer, G. J. (2007). A method to optimize the response range while maintaining Rorschach Comprehensive System validity. *Journal of Personality Assessment*, *89*, 149-161.
- Drew, L. R. H. (2005). Mortality and mental illness. *Australian and New Zealand Journal of Psychiatry*, *39*, 194-197.
- Dudek, S. Z., & Hall, W. B. (1984). Some test correlates of high level creativity in architects. *Journal of Personality Assessment*, *48*, 351-359.
- Eckblad, M. & Champman, L. J. (1983). Magical ideation as an indicator of schizotypy. *Journal of Consulting and Clinical Psychology*, *51*, 215-225.
- Elias, M. F., Elias, P. K., Sullivan, L. M., Wolf, P. A., & Agostino, R. B. D. (2005). Obesity, diabetes and cognitive deficit: The Framingham Heart Study. *Neurobiology of Aging*, *26S*, S11-S16.

- Endicott, N. A., & Endicott, J. (1964). Prediction of improvement in treated and untreated patients using the Rorschach Prognostic Rating Scale. *Journal of Consulting Psychology, 28*, 342-348.
- Exner, J. E. (1986). *The Rorschach: A Comprehensive System: Vol. 1. Basic foundations* (2<sup>nd</sup> ed.). New York: John Wiley & Sons.
- Exner, J. E. (2003). *The Rorschach: A Comprehensive System: Vol. 1. Basic foundations and principles of interpretation* (4<sup>th</sup> ed.). Hoboken, NJ: John Wiley & Sons.
- Fenchel, G. H., & Flapan, D. (1986). The developing ego in group psychotherapy. *Group, 10*, 195-210.
- Filmer-Bennett, G. (1955). The Rorschach as a means of predicting treatment outcome. *Journal of Consulting Psychology, 19*, 331-334.
- Freud, S. (1927). *The ego and the id*. London, UK: Hogarth Press (Original work published 1923).
- Freud, S. (1962). *Civilization and its discontents*. New York: Norton (Original work published 1930).
- Fuller, R. K., Lee, K. K., & Gordis, E. (1988). Validity of self-report in alcoholism research. *Alcoholism: Clinical and Experimental Research, 12*, 201-205.
- Gabbard, G. O., Coyne, L., Allen, J. G., Spohn, H., Colson, D. B., & Vary, M. (2000). Evaluation of intensive inpatient treatment of patients with severe personality disorders. *Psychiatric Services, 51*, 893-898.
- Garland, M., Hunter, D. J., Colditz, G. A., Spiegelman, D. L., Manson, J. E., Stampfer, M. J., et al. (1999). Alcohol consumption in relation to breast cancer risk in a

- cohort of United States women 25 - 42 years of age. *Cancer Epidemiology, Biomarkers, and Prevention*, 8, 1017-1021.
- Garson, G. D. (2008). Cox regression. *Statnotes: Topics in multivariate analysis*. Retrieved June 10<sup>th</sup>, 2009, from <http://faculty.chass.ncsu.edu/garson/PA765/cox.htm>
- Gelber, A. C., Hochberg, M. C., Mead, L. A., Wang, N. Y., Wigley, F. M., & Klag, M. J. (2000). Joint injury in young adults and risk of subsequent knee and hip osteoarthritis. *Annals of Internal Medicine*, 133, 321-328.
- Gottesman, I. I. (1959). More construct validation of the Ego-Strength Scale. *Journal of Consulting Psychology*, 23, 342-346.
- Graham, J. G. (1987). *The MMPI: A practical guide* (2<sup>nd</sup> ed.). New York: Oxford University Press.
- Graves, P. K., Mead, L. A., & Pearson, T. A. (1986). The Rorschach Interaction Scale as a potential predictor of cancer. *Psychosomatic Medicine*, 48, 549-563.
- Graves, P. L., Mead, L. A., Wang, N. Y., Liang, K. Y., & Klag, M. J. (1994). Temperament as a potential predictor of mortality: Evidence from a 41-year prospective study. *Journal of Behavioral Medicine*, 17, 111-126.
- Graves, P. L., & Thomas, C. B. (1981). Themes of interaction in medical students' Rorschach responses as predictors of midlife health and disease. *Psychosomatic Medicine*, 43, 215-225.
- Greenfield, N. S., Roessler, R., & Crosley, A. P. (1959). Ego strength and length of recovery from infectious mononucleosis. *Journal of Nervous and Mental Disease*, 128, 125-128.



- Greenwald, D. F. (1990). An external construct validity study of Rorschach personality variables. *Journal of Personality Assessment*, 55, 768-780.
- Greenwald, D. F., Harder, D. W., Ritzler, B. A., Strauss, J. S., & Kokes, R. F. (1988). The Last-Weiss Rorschach Ego-Strength Scale as a prognostic measure for psychiatric outpatients. *Perceptual and Motor Skills*, 67, 199-207.
- Grieg, T. C., Bell, M. D., Kaplan, E., & Bryson, G. (2000). Object relations and reality testing in early and late-onset schizophrenia. *Journal of Clinical Psychology*, 56, 505-517.
- Gross, C. P., Mead, L. A., Ford, D. E., & Klag, M. J. (2000). Physician, heal thyself? Regular source of care and use of preventative health services among physicians. *Archives of Internal Medicine*, 160, 3209-3214.
- Haller, N. (1982). *The reliability of Rorschach depressive indices in major depressive disorder*. Unpublished doctoral dissertation, United States International University, San Diego, CA.
- Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, 6, 278-296.
- Harder, D. W., Greenwald, D.F., Ritzler, B. A., Strauss, J. S., & Kokes, R. F. (1988). The Last-Weiss Rorschach Ego-Strength Scale as a prognostic measure for psychiatric inpatients. *Journal of Personality Assessment*, 52, 106-115.
- Harrower, M. R., & Steiner, M. E. (1951). *Large scale Rorschach techniques* (2<sup>nd</sup> ed.). Springfield, IL: Thomas.
- Hartmann, H. (1939). Ego psychology and the problem of adaptation. *International Journal of Psychoanalysis*, 21, 214-216.

- Hartmann, H. (1953). Contribution to the metapsychology of schizophrenia. *Psychoanalytic Study of the Child*, 8, 177-198.
- Hess, D. R. (2004). Retrospective studies and chart reviews. *Respiratory Care*, 49, 1171-1174.
- Hiroeh, U., Appleby, L., Mortensen, P. B., & Dunn, G. (2001). Death by homicide, suicide, and other unnatural causes in people with mental illness: A population-based study. *Lancet*, 358, 2110-2112.
- Hirschfeld, R. M. A. (1999). Personality disorders and depression: Comorbidity. *Depression and Anxiety*, 10, 142-146.
- Holiday, M. & Glidewell, R. (2000). The Bell Object Relations and Reality Testing Inventory (BORRTI). *Journal of Psychoeducational Assessment*, 18, 68-71.
- Howell, D. C. (2002). *Statistical Methods for Psychology* (5<sup>th</sup> ed.). Pacific Grove, CA: Duxbury.
- Hubert, H. B., Feinleib, M., McNamara, P. M., & Castelli, W. P. (1983). Obesity as an independent risk factor for cardiovascular disease: A 26-year follow-up of participants in the the Framingham Heart Study. *Circulation*, 67, 968-977.
- Hunter, C. G., & Goodstein, L. D. (1967). Ego strength and types of defensive and coping behavior. *Journal of Consulting Psychology*, 31, 432.
- James, W. (1879). Are we automata? *Mind*, 4, 1 – 22.
- Janson, H, & Olsson, U. (2001). A measure of agreement for interval and nominal multivariate observation. *Educational and Psychological Measurement*, 61, 277-289.

- Johnson, J. V., Hall, E. M., Ford, D. E., Mead, L. A., Levine, D. M., Wang, N. Y., et al. (1995). The psychosocial work environment of physicians: The impact of demands and resources of job dissatisfaction and psychiatric distress in a longitudinal study of Johns Hopkins Medical School graduates. *Journal of Occupational and Environmental Medicine, 37*, 1151-1159.
- Juni, S., & Stack, J. E. (2005). Ego function as a correlate of addiction. *American Journal of Addictions, 14*, 83-93.
- Kaneko, S. Y. (1984). The family system, temporary or permanent auxiliary ego: The differentiated fate of self-objects. *Clinical Social Work Journal, 12*, 303-319.
- Kaplan, E., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*, 457-481.
- Karasek, R. A. (1985). *Job Content Questionnaire and User's Guide*. Lowell: University of Massachusetts, Department of Work Environment.
- Kawachi, I., Sparrow, D., Spiro, A., Vokonas, P., Weiss, S. T. (1996). A prospective study of anger and coronary heart disease. *Circulation, 94*, 2090-2095.
- Kernberg, O. K. (1986). *Severe personality disorders: Psychotherapeutic strategies*. New Haven, CT: Yale University Press.
- King, G. K., & Schiller, M. (1960). Ego strength and type of defensive behavior. *Journal of Consulting Psychology, 24*, 215-217.
- Klag, M. J., Ford, D. E., Mead, L. A., He, J., Whelton, P. K., Liang, K. Y., et al. (1993). Serum cholesterol in young men and subsequent cardiovascular disease. *New England Journal of Medicine, 328*, 313-318.

- Klag, M. J., Wang, N. Y., Meoni, L. A., Brancati, F. L., Cooper, L.A., Liang, K. Y., et al. (2002). Coffee intake and risk of hypertension. *Archives of Internal Medicine*, 162, 657-662.
- Klopfer, B., Ainsworth, M. D., Klopfer, W. G., & Holt, R. R. (1954). *Developments in the Rorschach technique: Technique and theory*. New York, World Book Co.
- LaCroix, A. Z., Mead, L. A., Liang, K. Y., Thomas, C. B., & Pearson, T. A. (1986). Coffee consumption and coronary heart disease. *New England Journal of Medicine*, 315, 977-982.
- Laden, F., Neas, L. M., Tolbert, P. E., Holmes, M. D., Hankinson, S. E., Spiegelman, D. et al. (2000). Electric blanket use and breast cancer in the Nurses' Health Study. *American Journal of Epidemiology*, 152, 41-49.
- Last, U., & Weiss, A. A. (1976). Evaluation of ego strength based on certain Rorschach variables. *Journal of Personality Assessment*, 40, 57-66.
- Levine, D., & Cohen, J. (1962). Symptoms of ego strength measures as predictors of the outcome of hospitalization in functional psychoses. *Journal of Consulting Psychology*, 26, 246-250.
- Lichtenstein, R. (1984). Measuring job satisfaction of physicians in organized settings. *Medical Care*, 22, 56-68.
- Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: SAGE Publications.
- Loevinger, J. (1976). The meaning and measurement of ego development. *American Psychologist*, 21, 195-206.
- Loevinger, J., & Wessler, R. (1970). *Measuring ego development I: Construction and use of a sentence completion test*. San Francisco, CA: Jossey-Bass.

- Magnavita, J. J. (1997). *Restructuring personality disorders: A short-term dynamic approach*. New York, NY: Guilford.
- Mahler, M. S. (1952). On child psychosis and schizophrenia: Autistic and symbiotic infantile psychoses. *Psychoanalytic Study of the Child*, 7, 286-305.
- Manners, J., & Durkin, K. (2001). A critical review of the validity of ego development theory and its measurement. *Journal of Personality Assessment*, 77, 541-567.
- Mattis, S. (1973). *Dementia Rating Scale*. Psychological Assessment Resources, Inc., Odessa, FL.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1, 30-46.
- McWilliams, N. (1994). *Psychoanalytic diagnosis: Understanding personality structure in the clinical process*. New York, NY: Guilford.
- Molish, H. B., Molish, E. E., & Thomas, C. B. (1950). A Rorschach study of a group of medical students. *Psychiatric Quarterly*, 24, 744-774.
- Rollman, B., Mead, L.A., Wang, & Klag, M. (1997). Medical specialty and the incidence of divorce. *New England Journal of Medicine*, 336, 800-803.
- McHorney, C. A. (1996). Measuring and monitoring general health status in elderly persons: Practical and methodological issues in using the SF-36 Health Survey. *The Gerontologist*, 36, 571-584.
- Meyer, G. J., Erdberg, P., & Shaffer, T. W. (2007). Toward international normative reference data for the Comprehensive System. *Journal of Personality Assessment*, 89, 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., & Kurtz, J. E. (2006). Advancing personality assessment terminology: Time to retire “objective” and “projective” as personality test descriptors. *Journal of Personality Assessment, 87*, 223-225.
- Mihura, J. L., & Weinle, C. A. (2002). Rorschach training: Doctoral students' experiences and preferences. *Journal of Personality Assessment, 79*, 39-52.
- Molish, H. B., Molish, E. E., & Thomas, C. B. (1950). A Rorschach study of a group of medical students. *Psychiatric Quarterly, 24*, 744-774.
- Monte, C. F. (1987). *Beneath the mask: An introduction to theories of personality* (3<sup>rd</sup> ed.). New York: Holt, Rinehart, & Winston.
- Moore, A. A., Giuli, L., Gould, R., Hu, P., Zhou, K., Reuben, D., et al. (2006). Alcohol use, comorbidity, and mortality. *Journal of the American Geriatric Society, 54*, 757-762.
- Moore, A. A., Gould, R., Reuben, D. B., Greendale, G. A., Carter, K., Zhou, K., et al. (2005). Longitudinal patterns and predictors of alcohol consumption in the United States. *American Journal of Public Health, 95*, 458-465.
- Morgan, L., & Viglione, D. J. (1992). Sexual disturbances, Rorschach sexual responses, and mediating factors. *Psychological Assessment, 4*, 530-536.
- Naglieri, J. A., LeBuffe, P. A., & Pfeiffer, S. I. (1994). *Devereux Scales of mental Disorders manual*. San Antonio, TX: psychological Corporation.

- Norusis, M. J. (2008). *SPSS 16.0 Advanced Statistical Procedures Companion*. Upper Saddle River: Prentice Hall.
- Nunberg, H. (1931). The synthetic function of the ego. *International Journal of Psycho-Analysis*, *12*, 123-140.
- Nunnally, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill.
- Oh, K., Hu, F. B., Cho, E., Rexrode, K. M., Stampfer, M. J., Manson, J. E., et al. (2005). Carbohydrate intake, glycemic index, glycemic load, and dietary fiber in relation to risk of stroke in women. *American Journal of Epidemiology*, *161*, 161-169.
- O'Neill, R. M., & Bornstein, R. F. (1990). Oral-dependence and gender: Factors in help-seeking response set and self-reported psychopathology in psychiatric inpatients. *Journal of Personality Assessment*, *55*, 28-40.
- Onishi, M., Gjerde, P. F., & Block, J. (2001). Personality implications of romantic attachment patterns in young adults: A multi-method, multi-informant study. *Personality and Social Psychology Bulletin*, *27*, 1097-1110.
- Patton, M. J., Connor, G. E., & Scott, K. J. (1982). Kohut's psychology of self: Theory and measures of counseling outcome. *Journal of Counseling Psychology*, *29*, 268-282.
- Perry, W. (1994). Weiner's *Psychodiagnosis in Schizophrenia*: A classic revisited. *Journal of Personality Assessment*, *62*, 180-186.
- Perry, W. (2001). Incremental validity of the Ego Impairment Index: A re-examination of Dawes (1999). *Psychological Assessment*, *13*, 403-407.
- Perry, W., & Braff, D. L. (1994). Information-processing deficits and thought disorder in schizophrenia. *American Journal of Psychiatry*, *151*, 363-367.

- Perry, W., McDougall, A., & Viglione, D. (1995). A five-year follow-up on the temporal stability of the Ego Impairment Index. *Journal of Personality Assessment, 64*, 112-118.
- Perry, W., Minassian, A., Cadenhead, K., Sprock, J., & Braff, D. (2003). The use of the Ego Impairment Index across the schizophrenia spectrum. *Journal of Personality Assessment, 80*, 50-57.
- Perry, W., Moore, D., & Braff, D. (1995). Gender differences on thought disturbance measures among schizophrenic patients. *American Journal of Psychiatry, 152*, 1298-1301.
- 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.
- Perry, W., Viglione, D. J., & Braff, D. L. (1992). The Ego Impairment Index and schizophrenia: A validation study. *Journal of Personality Assessment, 59*, 165-175.
- Quay, H. C. (1963). Ego-strength and psychiatric diagnosis. *Psychological Reports, 13*, 170.
- Quinn, R. P., & Staines, G. L. (1979). *The 1977 quality of employment survey*. Ann Arbor, MI: University of Michigan, Division for Social Research.
- Ritzler, B., & Nalesnick, D. (1990). The effect of inquiry on the Exner Comprehensive System. *Journal of Personality Assessment, 55*, 647-656.
- Roberts, B. W., Kuncel, N., Shiner, R., N., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socio-



- economic status, and cognitive ability for predicting important life outcomes. *Perspectives in Psychological Science*, 2, 313-345.
- Rollman, B. L., Mead, L. A., Wang, N. Y., & Klag, M. J. (1997). Medical specialty and the incidence of divorce. *New England Journal of Medicine*, 336, 800-803.
- Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin*, 94, 18-38.
- Sadowski, C. J., Woodward, H. R., Davis, S. F., & Elsbury, D. L. (1983). Sex differences in adjustment correlates of locus of control dimensions. *Journal of Personality Assessment*, 47, 627-631.
- Schill, T., & Thomsen, D. (1987). Anger, hostility, and the Barron Ego Strength Scale. *Psychological Reports*, 60, 1113-1114.
- Shaffer, J. W., Duszynski, K. R., & Thomas, C. B. (1981). Orthogonal dimensions of individual and group forms of the Rorschach. *Journal of Personality Assessment*, 15, 230-239.
- Shaffer, J. W., Duszynski, K. R., & Thomas, C. B. (1982). Family attitudes in youth as a possible precursor of cancer among physicians: A search for explanatory mechanisms. *Journal of Behavioral Medicine*, 5, 143-163.
- Shedler, J., & Block, J. (1990). Adolescent drug use and psychological health: A longitudinal inquiry. *American Psychologist*, 45, 612-630.
- Shybut, J. (1970). Internal vs. external control, time perspective and delay of gratification of high and low ego strength groups. *Journal of Clinical Psychology*, 26, 430-431.

- Sinnett, E. R. (1962). The relationship between the Ego Strength Scale and rated in-hospital improvement. *Journal of Clinical Psychology, 18*, 46-47.
- Sommers, M. S., Dyehouse, J. M., Howe, S. R., Lemmink, J., Volz, T., Manharth, M., et al. (2000). Validity of self-reported alcohol consumption in nondependent drinkers with unintentional injuries. *Alcoholism: Clinical and Experimental Research, 24*, 1406-1413.
- SPSS, Inc. (2007). *SPSS 16.0 brief guide*. Chicago, IL: Author.
- Stewart, A. L., Greenfield, S., Hays, R. D., Wells, K., Rogers, W. H., Berry, et al. (1989). Functional status and well-being of patients with chronic conditions: Results from the Medical Outcome Study. *Journal of the American Medical Association, 262*, 907-913.
- Stokes, J. M., Pogge, D. L., Powell-Lunder, J., Ward, A. W., Bilgner, 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.
- Sutton, P. M., & Swenson, C. H. (1983). The reliability and concurrent validity of alternative methods for assessing ego development. *Journal of Personality Assessment, 47*, 468-475.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4<sup>th</sup> ed.). Needham Heights, MA: Allyn & Bacon.
- Thomas, C. B. (1971). Suicide among us: II. Habits of nervous tension as potential predictors. *Johns Hopkins Medical Journal, 129*, 190-201.

- Thomas, C. B. (1976). Precursors of premature disease and death: The predictive potential of habits and family attitudes. *Annals of Internal Medicine*, 85, 653-658.
- Thomas, C. B., & Duszynski, K. R. (1985). Are words of the Rorschach predictors of disease and death? The case of "whirling." *Psychosomatic Medicine*, 47, 201-211.
- Thomas, C. B., Ross, D. C., Brown, B. S., & Duszynski, K. R. (1973). A prospective study of the Rorschachs of suicides: The predictive potential of pathological content. *Johns Hopkins Medical Journal*, 132, 334-360.
- Thomas, C. B., Santora, P. B., & Shaffer, J. W. (1980). Health of physicians in midlife in relation to use of alcohol. *Johns Hopkins Medical Journal*, 146, 1-10.
- Torre, D. M., Wang, N. Y., Meoni, L. A., Young, J. H., Klag, M. J., & Ford, D. E. (2005). Suicide compared to other causes of mortality in physicians. *Suicide and Life Threatening Behavior*, 35, 146-153.
- Vaillant, G. E. (1977). *Adaptation to life*. Boston, MA: Little-Brown.
- Vaillant, G. E. (1983). *The natural history of alcoholism*. Cambridge, MA: Harvard University Press.
- Vaillant, G. E. (1995). *Natural history of alcoholism revisited*. Cambridge, MA: Harvard University Press.
- Vaillant, G. E. (2002). *Aging well*. Boston, MA: Little-Brown.
- Vaillant, G. E. (1994). Ego mechanisms of defense and personality psychopathology. *Journal of Abnormal Psychology*, 103, 44-50.
- Vaillant, G. E., & Mukamal, K. (2001). Successful aging. *American Journal of Psychiatry*, 158, 839-847.

- Vaillant, G. E. & Vaillant, C. O. (1990). Natural history of male psychological health, XII. *American Journal of Psychiatry*, 147, 31-37.
- Van den Berg, M. D., Oldehinkel, A. J., Bouhuys, A. L., Brilman, E. I., Beekman, A. T. F., Ormel, J. (2001). Depression in later life: Three etiologically different subgroups. *Journal of Affective Disorders*, 65, 19-26.
- Viglione, D. J. (1990). Severe disturbance or trauma-induced adaptive reaction: A Rorschach child case study. *Journal of Personality Assessment*, 55, 280-295.
- Viglione, D. J., Perry, W., Jansak, D., Meyer, G., & Exner, J. E. (2003). Modifying the Rorschach Human Experience Variable to create the Human Representational Variable. *Journal of Personality Assessment*, 81, 64-73.
- Viglione, D. J., Perry, W., Meyer, G. (2003). Refinements in the Rorschach Ego Impairment Index incorporating the Human representational Variable. *Journal of Personality Assessment*, 81, 149-156.
- Wadsworth, M. E. J., Butterworth, S. L., Hardy, R. J., Kuh, D. J., Richards, M., Langenberg, C., et al. (2003). The life course prospective design: An example of benefits and problems associated with study longevity. *Social Science and Medicine*, 57, 2193-2205.
- Wang, N. Y., Young, J. H., Ford, D. E., Meoni, L. A., Erlinger, T., & Klag, M. J. (in press). Blood pressure change and risk of hypertension associated with parental hypertension: The Johns Hopkins Precursors Study. *Archives of Internal Medicine*.
- Weiner, I. P. (1966). *Psychodiagnosis in schizophrenia*. New York, NY: Wiley.

- Werner, H. (1948). *The comparative psychology of mental development* (Revised ed.).  
Chicago, IL: Follet.
- Whalley, L. J., & Deary, I. J. (2001). Longitudinal cohort study of childhood IQ and  
survival up to age 76. *British Medical Journal*, 322, 1-5.
- Worden, J. T., & Sobel, H. J. (1978). Ego strength and psychosocial adaptation to cancer.  
*Psychosomatic Medicine*, 40, 585-59.