JOURNAL OF PERSONALITY ASSESSMENT
2016, VOL. 98, NO. 3, 289–297
http://dx.doi.org/10.1080/00223891.2015.1107573

Routledge
Taylor & Francis Group

Convergent Validity of the Early Memory Index in Two Primary Care Samples

John H. Porcerelli,[1] Rosemary Cogan,[2] Katherine A. Melchior,[3] Matthew J. Jasinski,[4] Laura Richardson,[3] Shannon Fowler,[3] Pierre Morris,[1] and William Murdoch[1]

[1]Department of Family Medicine & Public Health Sciences, Wayne State University School of Medicine; [2]Department of Psychology, Texas Tech University; [3]Department of Psychology, University of Detroit Mercy; [4]Department of Psychology, Wayne State University

ABSTRACT

Karliner, Westrich, Shedler, and Mayman (1996) developed the Early Memory Index (EMI) to assess mental health, narrative coherence, and traumatic experiences in reports of early memories. We assessed the convergent validity of EMI scales with data from 103 women from an urban primary care clinic (Study 1) and data from 48 women and 24 men from a suburban primary care clinic (Study 2). Patients provided early memory narratives and completed self-report measures of psychopathology, trauma, and health care utilization. In both studies, lower scores on the Mental Health scale and higher scores on the Traumatic Experiences scale were related to higher scores on measures of psychopathology and childhood trauma. Less consistent associations were found between the Mental Health and Traumatic Experiences scores and measures of health care utilization. The Narrative Coherence scale showed inconsistent relationships across measures in both samples. In analyses assessing the overall fit between hypothesized and actual correlations between EMI scores and measures of psychopathology, severity of trauma symptoms, and health care utilization, the Mental Health scale of the EMI demonstrated stronger convergent validity than the EMI Traumatic Experiences scale. The results provide support for the convergent validity of the Mental Health scale of the EMI.

ARTICLE HISTORY
Received 29 May 2014
Revised 28 August 2015

Early memories have been of interest to scholars throughout history (e.g., Carruthers, 1966). Research involving early memories was reported as early as 1895 by Miles, who collected responses of 100 women at Wellesley College to questions that included one about their earliest memory. Freud (1899/1961) wrote of the earliest memories of childhood and described them as screen memories that result from the same mental processes as those leading to neurotic symptoms. Freud noted an 1897 study by V. Henri and C. Henri that explored the content of early recollections from 123 normal adults. In a paper written from an Adlerian perspective, Schrecker (1913/1973) took the view that a person "can recollect only what is somehow related to his present situation" (p. 148).

Since these early investigations, several formal systems for assessing early memories have been developed. For example, Langs and his colleagues (Langs, 1965a, 1965b; Langs, Rothenberg, Fishman, & Reiser, 1960) developed the *Manual for the Scoring of the Manifest Content of the Earliest Memory* (Langs et al., 1960) and found that the content of early memories successfully differentiated hospitalized schizophrenic and schizoid patients from people hospitalized with hysterical character disorders. Bruhn (1985, 1990) developed a cognitive-perceptual interpretive method for earliest memories, which later became known as the Comprehensive Early Memories Scoring System. Drawing from ego psychological and cognitive-perceptual theories, Bruhn (1985) hypothesized that a person's needs, desires, fears, and core beliefs influence the perceptual process as well as the recall of events. From this perspective, early memories reveal important aspects of personality structure.

Mayman and Faris (1960) reported work on early memories and developed a system for coding the libidinal level of self–other representations in early memories. From an object relations and ego psychological approach, Mayman (1963, 1968) hypothesized that early memories are selected unconsciously to reveal enduring self–other images and core relational themes and conflicts. Shedler, Mayman, and Manis (1993) extended Mayman's conceptualizations about early memories to address the issue of illusory mental health. Participants who reported positive mental health on self-report scales, but were viewed by four experienced psychodynamic clinicians as distressed based on the transcripts of their early memories, were placed in the illusory mental health group. This group displayed greater average and maximum cardiac reactivity to stressors than a genuine (positive) mental health group ($d = .76$ and 1.03, respectively) and a genuine distressed group (who showed consistency between self-report and early memory ratings of distress, $d = .85$ and 1.14, respectively). These findings suggested that the denial of distress by the illusory mental health group is a risk factor for developing cardiac illness. Attempts to replicate these findings have had mixed results. Taylor, Lerner, Sherman, Sage, and McDowell (2003) failed to replicate these findings in a college population working with Early Memory Index (EMI)

CONTACT John H. Porcerelli jporcer@med.wayne.edu Department of Family Medicine & Public Health Sciences, Wayne State University School of Medicine, 1135 W. University Drive, Suite 250, Rochester Hills, MI 48307.

ratings of distress by a single experienced clinician; however, Nehrig (2014) did replicate these findings in a diverse sample of college students.

One limitation of the original illusory mental health study was the lack of a formal system for coding early memories. To address this concern, Karliner, Westrich, Shedler, and Mayman (1996) developed the Adelphi Early Memory Index (AEMI), referred to as the Early Memory Index in a subsequent article by Shedler, Karliner, and Katz (2003). Based on a factor analysis of their initial item pool, three factors emerged: Mental Health,[1] Narrative Coherence, and Traumatic Experiences. Although the Narrative Coherence and Traumatic Experiences factors correlated significantly with composite clinical judgments of mental health, in a stepwise regression analysis, neither factor added to the prediction of mental health above and beyond the Mental Health factor.

Cousineau and Shedler (2006) used the Mental Health scale to predict the number of times students visited a university-based health clinic, the number of verified illness-related visits to a university health clinic within the past year, and retrospective reports of health care utilization during the academic year. The Mental Health scale was related to the actual number of visits to the clinic and clinically verified illness in the previous year (r = –.32 and –.33, respectively). In the development of the Early Spiritual Memory Test (ESMT), Lin (2006) found that the Mental Health score from the EMI did not significantly correlate with the Mental Health score from the ESMT.

In theory, patients who have difficulty telling a coherent story might not have fully processed distressing feelings associated with painful or traumatic memories, or their recall of the memories might be colored by defensive processes (Karliner et al., 1996). Foa, Molnar, and Cashman (1995) and Van Minnen, Wessel, Dijkstra, and Roelofs (2002) demonstrated that narratives told at the end of exposure therapy for posttraumatic stress disorder (PTSD) became more coherent when compared to narratives at the beginning of treatment. Thus, the Narrative Coherence scale could contribute to the assessment of trauma as well as overall mental health. Similarly, the Traumatic Experiences scale of the EMI could provide important information, especially to assessments in which the referral questions are not related directly to trauma. Because the EMI does not specifically ask for memories of trauma, any such memories can alert clinicians to obtain a more detailed assessment of trauma. In two studies, we explore the associations between all three EMI scales and self-report measures of psychopathology (depression and borderline personality disorder symptoms). These disorders were chosen because the prevalence of major depressive disorder in primary care ranges from 10% to 14% (Katon & Schulberg, 1992), and the prevalence of borderline personality disorder is four times greater in primary care than in community samples (Gross et al., 2002). Borderline personality disorder also poses the greatest threat to the doctor–patient relationship in primary care (Porcerelli & Huprich, 2007). Therefore, the assessment of early memories in a primary care setting can provide useful information about the level of pathology experienced by a patient that the patient himself or herself might not be able to report directly.

To date, of the EMI subscales, only the Mental Health scale has been studied outside of Shedler's lab. Thus, we assessed the validity of all three EMI factor scales using criterion measures that are in current use in each primary care setting and have demonstrated clinical utility, validity, or both in such settings. We included measures of health care utilization to replicate the earlier findings of Cousineau and Shedler (2006) described earlier. Due to the lack of validity data for the EMI, magnitudes of effect sizes for the EMI and criterion variables are speculative.

We hypothesized the following:

1. EMI Mental Health and Narrative Coherence will negatively and significantly correlate with measures of depression, borderline personality disorder symptoms, and severity of trauma symptoms. We expect the magnitude of these correlations to approach a medium effect size (r = .30) according to Cohen (1988). We expect that the correlations between EMI Mental Health and health care utilization to be negative and approach a small to medium effect size (r = .20) for emergency room (ER) visits and a small effect size (r = .15) for hospitalizations. We expect the ER visit correlation to be slightly larger in magnitude because mental health problems might cause a person to go to an ER, but hospitalizations require more objective physical pathology.
2. EMI Traumatic Experiences will positively and significantly correlate with measures of depression and borderline personality disorder symptoms. We expect these correlations to approach a medium effect size (r = .30). We expect EMI Traumatic Symptoms will positively and significantly correlate with severity of trauma symptoms and approach a medium to large effect size (r = .40). We expect that the correlations between EMI Traumatic Experiences and health care utilization will be positively correlated and approach a small to medium effect size (r = .20).

In addition to assessing construct validity through the analysis of individual correlations, construct validity will also be assessed through a methodology developed by Westen and Rosenthal (2003), called $r_{contrast\text{-}CV}$. This methodology provides a single metric for testing construct validity across several correlations. The methodology essentially assesses the fit between predicted and actual correlations between a given construct (e.g., mental health) and a group of variables used to assess the validity of the construct. The calculation takes into account sample size and the size of the intercorrelations among criterion variables. The calculation of $r_{contrast\text{-}CV}$ can be found in Appendix B in Westen and Rosenthal (2003).

Study 1: Urban primary care

Participants

As part of an effort to understand women's health (Porcerelli, Cogan, Markova, Murdoch, & Porcerelli, 2010), we recruited women from the waiting room of a university-based family medicine residency training clinic in a large Midwestern city. A

[1] Although Shedler, Karliner, and Katz (2003) named the first EMI scale Mental Health/Distress, they referred to the scale as "mental health" in the text of later papers (e.g., Cousineau & Shedler, 2006). We believe Mental Health is a less confusing term and use this throughout in referring to the first EMI scale.

total of 161 Medicaid-insured or Medicaid-eligible women (a requirement of a Medicaid-funded study) who consecutively presented to the clinic were asked to participate. Interested patients returned to the clinic to complete self-report measures and an interview that included the early memory interview procedure. Of 110 women who agreed to participate in the study, 103 had viable video recordings (i.e., the recordings of 7 were corrupted and could not be used) that included the early memory interview. The average age of the women was 35.25 years ($SD = 9.61$). Forty-nine (48%) were single, and 33 (32%) were separated or divorced; 91 (88%) were African American, 87 (84%) had household incomes of $20,000 or below, and 77 (75%) had 12 or more years of education. Participants were provided an $80 honorarium for their participation. The study was approved by the institutional review boards of Wayne State University and the Michigan Department of Community Health.

Procedures

After patients were informed about the study and consent was received, participants completed all study questionnaires prior to providing early memory narratives. Questionnaires were obtained and early memories were collected by a doctoral- and master's-level psychologist and doctoral-level social workers. These clinicians received training in administering the early memory protocol. Patients provided 10 early memories, 8 of which were recommended by Fowler, Hilsenroth, and Handler (1995; earliest, next earliest, earliest of mother, earliest of father, earliest of school, earliest of feeling warm and snug, earliest of eating or being fed, and earliest of a special object), plus two others (happiest and saddest early memory). The early memory interview began by asking patients to take a moment to relax, take a deep breath, and to let their thoughts go back to childhood. Then they were asked to "think back as far as you can and try to recall your earliest memory." Patients who did not report a specific memory were prompted to report a specific incident or memory. At the end of each memory, patients were asked to give their impressions of self and others in the memory and the mood or feeling of the memory. The early memory interviews were videotaped, and scores were coded directly from the videotapes. Early memories were coded for all three scales of the EMI: Mental Health, Narrative Coherence, and Traumatic Experiences. Twenty-six (25%) videotapes were double coded for interrater reliability by a doctoral-level psychologist and psychology doctoral student, both of whom received extensive training in coding early memories. The coders were blind to all health-related information about the patients. The remaining protocols were coded by the doctoral student.

Measures

Early memories

The EMI (Shedler et al., 2003) is comprised of three factors: Mental Health, Narrative Coherence, and Traumatic Experiences. Each item on each scale is evaluated on a 5-point Likert-type scale, ranging from 1 (*not applicable*) to 5 (*highly applicable*). There are nine items in the Mental Health scale; five are positively worded (e.g., "Predominant affect tone is positive"), and four are negative (e.g., "Others are depicted as malevolent") and are reverse scored so that higher scores on the scale indicate better mental health. There are five items on the Narrative Coherence scale. Two are positively worded (e.g., "The memories seem real, full bodied, palpable, easy to imagine"), and three are negatively worded (e.g., "The memories are lacking in evocative detail") and are reverse scored so that higher scores on the scale indicate greater narrative coherence. There are four items on the Traumatic Experiences scale. All four are negative (e.g., "Subject experiences others as deliberately inflicting physical injury on him or her"), and higher scores thus represent greater trauma.

Depression

The Patient Health Questionnaire–9 (PHQ–9; Kroenke, Spitzer, & Williams, 2001) is a 9-item self-report scale, based on *Diagnostic and Statistical Manual of Mental Disorders* (4th ed. [*DSM–IV*]; American Psychiatric Association, 2000) criteria for major depressive disorder, developed for use in primary care patients. Items are rated on a 4-point Likert-type scale ("over the last 2 weeks") ranging from 0 (*not at all*) to 3 (*nearly every day*). Scores range from 0 to 27, with a score of 10 or more indicating moderate to severe depression. The validity of the PHQ–9 has been supported in studies from primary care (Spitzer, Kroenke, & Williams, 1999) and obstetric-gynecologic practices (Spitzer, Williams, Kroenke, Hornyak, & McMurray, 2000). The internal consistency of the PHQ–9 was good for this study ($\alpha = .89$). Total scores were used to indicate symptom severity.

Borderline personality disorder

The Borderline Personality Disorder (BPD) module of the Personality Diagnostic Questionnaire (PDQ–4+; Hyler, 1994) is a 9-item true–false self-report scale developed to reflect *DSM–IV* criteria for BPD. The time frame used for the ratings is "over the past several weeks." Scores of five or more are assigned a diagnosis (Hyler, 1994). Okada and Oltmanns (2009) reported adequate test–retest reliability and convergent validity for the full PDQ–4+, and Davison, Leese, and Taylor (2001) indicated that the full scale also had adequate screening properties for identifying the presence or absence of personality disorders. Total scores were used to indicate symptom severity and demonstrated an internal consistency of $\alpha = .67$ for this study.

Childhood trauma

The Childhood Trauma Questionnaire–Short Form (CTQ–SF; Bernstein & Fink, 1998) is a 28-item self-report scale that measures childhood physical, sexual, and emotional abuse, and emotional and physical neglect. Each subscale includes five items, and each item is rated on a 5-point Likert-type scale ranging from 1 (*never true*) to 5 (*very often true*). Cutoff scores for moderate to severe abuse are 10 or more for physical abuse, 8 or more for sexual abuse, and 13 or more for emotional abuse. Cutoff scores for moderate to severe neglect are 15 or more for emotional neglect and 10 or more for physical neglect.

The reliability and validity of the CTQ–SF were reported by Bernstein et al. (2003). The internal consistency of the physical, sexual, and emotional abuse subscales in this study were

$\alpha = .89$, $\alpha = .95$, and $\alpha = .91$, respectively. The internal consistency of the emotional and physical neglect scales in this study was $\alpha = .87$ and $\alpha = .71$, respectively. The total of each of the scales was used in this study.

Posttraumatic stress disorder

The Posttraumatic Stress Disorder Checklist–Civilian version (PCL–C; Weathers, Litz, Herman, Huska, & Keane, 1993) is a 17-item self-report scale for assessing *DSM–IV* PTSD. Each item is rated on a 5-point Likert-type scale ("in the past month"), ranging from 1 (*not at all*) to 5 (*extremely*). Scores range from 17 to 85 with a cutoff score of 44 recommended for use with a civilian population (Ruggiero, Del Ben, Scotti, & Rabalais, 2003). The reliability and validity of the PCL–C have been reported by Ruggiero et al. (2003). Symptoms can be grouped according to PTSD clusters (intrusion, avoidance, and hyperarousal) or as a total score. Internal consistency of the PCL–C in this study was $\alpha = .94$, and total scores were used as an indicator of symptom severity.

Health care utilization

Patients reported the frequency of ER visits and hospitalizations over the past year on 6-point scales ranging from 0 (*0 visits*) to 5 (*5 or more visits*). These items were extracted from the Multidimensional Health Profile–Health Functioning scale (MHP–H; Karoly, Ruehlman, & Lanyon, 2005; Ruehlman, Lanyon, & Karoly, 1998).

Data analysis

We calculated the means, standard deviations, and internal consistency reliability of each EMI scale. Means and standard deviations for depression severity, BPD symptoms, child maltreatment severity, PTSD symptoms, number of ER visits, and number of overnight hospitalizations were also calculated. Ninety-six percent ($n = 99$) of the study protocols had complete self-report data. Missing data (1 or 2 items) occurred on just two symptom scales, although the totals of those scales were used for data analysis. We calculated Pearson correlations between the EMI scale scores and other variables. We used mean scores for the EMI scales for all analyses, following Shedler et al. (2003). There were no missing early memory narratives. Following the guidelines by Westen and Rosenthal (2003), we tested the construct validity of the three EMI scales. We ran $r_{contrast\text{-}CV}$ effect size correlations comparing the hypothesized and actual magnitudes of the three EMI scales with all outcome measures: psychopathology, severity of trauma symptoms, and health care utilization.

Results

The mean scores of the women on each of the measures are shown in Table 1. All of the variables were normally distributed (Curran, West, & Finch, 1996), with the exception of the EMI Traumatic Experiences scale, which was moderately skewed (skewness = 2.28). As a group, the women had mild levels of depression, had a moderate to severe level of childhood sexual abuse, and approached the cutoff score for moderate to severe childhood physical abuse. The mean scores were not elevated for BPD symptoms, PTSD symptoms, childhood emotional abuse, childhood emotional neglect, or childhood physical neglect.

Table 1. Descriptive statistics for the urban and suburban samples.

	Urban[a]		Suburban[b]	
Measures	*M*	*SD*	*M*	*SD*
EMI Mental Health	3.32	.32	3.33	.44
EMI Narrative Coherence	3.32	.62	3.65	.41
EMI Traumatic Experiences	1.17	.25	1.34	.28
Depression[c]	7.80	6.30	1.63	1.91
Borderline personality disorder symptoms	2.76	2.27	2.13	1.91
Childhood abuse and neglect				
Physical abuse	9.42	5.31	6.65	3.88
Sexual abuse	10.24	7.17	6.92	4.81
Emotional abuse	10.70	6.33	—	—
Emotional neglect	11.30	5.19	—	—
Physical neglect	7.95	3.77	—	—
Posttraumatic stress disorder symptoms[d]	34.92	14.37	1.25	1.43
Emergency room	1.21	1.31	1.18	1.28
Hospital overnight	.51	.96	.52	.89

Note. EMI = Early Memory Index.
[a] $n = 103$.
[b] $n = 72$.
[c] Depression was assessed in the urban sample with the Patient Health Questionnaire–9 (9 items, range = 0–27) and in the suburban sample with the Patient Health Questionnaire–2 (2 items, range = 0–6).
[d] Posttraumatic stress disorder was assessed in the urban sample with the Posttraumatic Stress Disorder Checklist–Civilian version (17 items, range = 17–85) and in the suburban sample with the Primary Care Post Traumatic Stress Disorder Questionnaire (4 items, range = 0–4).

Reliability and validity of the EMI

Mean scores for each EMI item as well as corrected item-to-total correlations are reported in Table 2. Adequate levels of internal consistency (Cronbach's α) were obtained for each of the EMI scales: Mental Health, .82; Narrative Coherence, .75; and Traumatic Experiences, .84. Interrater reliabilities (intraclass correlation coefficients [ICC], one-way random effect) with Spearman–Brown correction for double coding were Mental Health, .90; Narrative Coherence, .75; and Traumatic Experiences, .84. Interrater reliability reached the excellent range (ICC > .75; Shrout & Fleiss, 1979) for the three EMI scales.

Intercorrelations of the EMI scales were Mental Health and Narrative Coherence, $r = .49$, $p < .001$; Mental Health and Traumatic Experiences, $r = -.60$, $p < .001$; and Narrative Coherence and Traumatic Experiences, $r = -.06$, $p = .64$. The Mental Health and Traumatic Experiences scores significantly correlated in the expected directions with measures of psychopathology, trauma symptoms, and one measure of health care utilization (overnight hospitalizations), each with a medium effect size. Neither Mental Health nor Traumatic Experiences were significantly correlated with ER use. The Narrative Coherence scale was significantly correlated with BPD symptoms with a small to medium effect size (see Table 3).

Validity coefficients from the $r_{contrast\text{-}CV}$ analyses for the suburban sample are also reported in Table 3. The magnitude of the $r_{contrast\text{-}CV}$ coefficient in the urban sample indicates that the construct validity of the Mental Health score was excellent (Westen & Rosenthal, 2003), the $r_{contrast\text{-}CV}$ coefficient for

Table 2. Early Memory Index (EMI) item means and corrected item–total correlations.

	Urban			Suburban		
EMI Scale	*M*	*SD*	*r*	*M*	*SD*	*r*
Mental Health/Distress						
1. Predominant affect tone is positive (e.g., memories convey happiness, contentment, well-being, excitement, etc.).	2.88	.51	.81	2.76	.59	.63
2. Predominant affect tone is negative (e.g., memories convey sadness, anger, frustration, fear, hurt, etc.).[a]	3.67	.56	.86	3.69	.58	.68
3. The memories have predominantly positive outcomes (e.g., subject ultimately succeeds, experiences gratification, etc.).	2.92	.53	.72	2.87	.61	.63
4. The memories have predominantly negative outcomes (e.g., subject ultimately fails, experiences pain, frustration, etc.).[a]	3.94	.51	.83	3.77	.58	.54
5. Others are depicted as benevolent; they are seen as sources of gratification, pleasure, comfort, security, and so on.	2.88	.57	.73	2.94	.61	.45
6. Subject comes across as confident, self-assured.	3.55	.50	.46	2.65	.46	.52
7. Subject comes across as ignored, deprived, not cared for.[a]	3.71	.53	.69	4.34	.57	.56
8. Others are depicted as malevolent; they are seen as sources of frustration, pain, punishment, injury, and so on.[a]	4.14	.56	.79	3.89	.59	.64
9. Caregivers are portrayed as abandoning or underproductive.[a]	4.22	.48	.80	4.44	.59	.60
Narrative Coherence						
1. The memories seem "real," full-bodied, palpable, easy to imagine; one gets a sense of being there; one can empathize with subject's feelings.	2.53	.53	.89	2.91	.63	.64
2. The memories are lacking in evocative detail; descriptions seem thin, two-dimensional, clichéd, or otherwise lacking in supportive detail; one cannot really get a sense of what it was like to be there; one cannot readily empathize with subject's feelings.[a]	2.52	.77	.86	2.45	.97	.57
3. Emotions reported by the subject are congruent with the content of the memory; that is, details of the memory are consistent with the reported emotions.	3.55	.52	.70	3.46	.47	.56
4. Emotions reported by the subject are incongruent with the content of the memory; that is, supporting details are lacking, or details of the memory are inconsistent with the reported emotions.[a]	4.10	.52	.52	4.58	.61	.51
5. The narrative lacks inner narrative coherence; there are pieces that do not seem to "fit"; there is a sense that things have been omitted or distorted.[a]	3.37	.89	.86	3.60	.78	.73
Traumatic Experiences						
1. Subject experiences others as deliberately inflicting physical injury on him or her.	1.17	.31	.82	1.36	.44	.72
2. There is at least one memory of an extremely traumatic event falling outside the realm of normal childhood experience (e.g., of physical or sexual abuse, violence between parents, a parent drunk and out of control, etc.).	1.27	.45	.90	1.21	.47	.91
3. Subject perceives others as deliberately inflicting injury on others (e.g., domestic violence, schoolyard fights).	1.11	.27	.79	1.16	.26	.60
4. Subject experiences injury or illness.	1.15	.29	.52	1.66	.49	.86

[a]These items are reverse scored and have been reverse scored here.

Traumatic Experiences score was good, and the $r_{contrast\text{-}CV}$ coefficient for the Narrative Coherence score was poor.

Fisher *r*-to-*z* comparisons were calculated for correlations between Mental Health and Traumatic Experiences and measures of psychopathology, severity of trauma symptoms, and health care utilization. None of the comparisons was significantly different. Thus, the magnitude of the relationships between Mental Health and criterion variables were equal to the magnitude of relationships between Traumatic Experiences and criterion variables.

Study 2: Suburban primary care

Participants

We recruited women ($n = 48$) and men ($n = 24$) from the waiting room of a suburban university-based family medicine residency training clinic. One hundred and forty consecutive patients were asked to participate in a study of relationships and health. Patients in the study had a mean age of 34.29 years ($SD = 13.49$). Thirty (42%) were single, and 27 (38%) were married; 45 (63%) were White, and 18 (25%) were African American; 49 (68%) had household incomes of less than $60,000; and 66 (92%) had 12 or more years of education. Patients were provided a $10 honorarium for their participation. The study was approved by the institutional review board of Wayne State University.

Procedures

After patients were informed about the study and consent was received, patients completed self-report measures in the waiting room prior to their visit with their physician and completed the early memory interview in the exam room before or immediately after their appointment. In this study, briefer measures were used to reduce patient flow problems. Thus, a two-item version of the PHQ–9, a different (briefer) PTSD scale, and fewer early memory narratives were collected. Also, in this

Table 3. Early Memory Index (EMI) correlates in urban ($n = 103$) and suburban ($n = 72$) primary care samples.

	Psychopathology		Severity of Trauma Symptoms		
EMI Factor	Depression Severity[a]	Borderline Personality Disorder Symptoms	Childhod Sexual Abuse	Childhood Physical Abuse	Childhood Emotional Abuse
Mental health					
Urban	−.34***	−.39***	−.47***	−.40***	−.38***
Suburban	−.40***	−.39***	−.36**	−.25*	—
Narrative coherence					
Urban	−.02	−.22*	−.13	−.11	−.03
Suburban	−.33**	−.17	−.26*	−.09	—
Traumatic experiences					
Urban	.20*	.34***	.45***	.46***	.31***
Suburban	.47***	.38***	.17	.49***	—

Severity of Trauma Symptoms			Health Care Utilization		
Childhood Emotional Neglect	Childhood Physical Neglect	Posttraumatic Stress Disorder[b]	Emergency RoomVisits	Hospital Overnights	Construct Validity ($r_{contrast\text{-}cv}$)
−.48***	−.41***	−.29**	−.14	−.30**	(.84)
—	—	−.20	−.15	−.07	(.74)
.01	−.06	.01	.02	−.10	(.26)
—	—	−.06	−.09	.13	(.42)
.47***	.49***	.30**	.05	.24*	(.71)
—	—	.30**	.40***	.39***	(.50)

[a] Depression was assessed in the urban sample with the Patient Health Questionnaire–9 (9 items, range = 0–27) and in the suburban sample with the Patient Health Questionnaire–2 (2 items, range = 0–6).
[b] Posttraumatic stress disorder (PTSD) was assessed in the urban sample with the Posttraumatic Stress Disorder Checklist–Civilian version (17 items, range = 17–85) and in the suburban sample with the Primary Care Post Traumatic Stress Disorder Questionnaire (4 items, range = 0–4).
*$p < .05$.
**$p < .01$.
***$p < .001$.

study, patients completed self-report measures prior to entering the exam room, whereas in Study 1, patients who were interested in participation returned at a later date to complete all research materials. Patients provided five early memories (earliest, mother, father, high point, and low point) as reported by Cousineau and Shedler (2006). Early memories were elicited from patients by a trained doctoral student while they were in the exam room prior to their appointment with their primary care physician, described in Study 1. Memories were written verbatim by a research assistant and later transcribed. All memories were double coded by graduate research assistants blind to all other patient data.

Measures

The EMI, PDQ–4+, CTQ scales, and the scales for assessing ER visits and overnight hospitalizations were described in Study 1. The internal consistency of the BPD scale from the PDQ–4+ for this study was $\alpha = .68$. Internal consistency of the childhood Sexual and Physical Abuse scales of the CTQ were $\alpha = .95$ and $\alpha = .90$, respectively.

Depression

The Patient Health Questionnaire–2 (PHQ–2; Kroenke, Spitzer, & Williams, 2003), a 2-item depression screener developed for primary care practices, is made up of the first two items from the PHQ–9 (Kroenke et al., 2001): *Little interest or pleasure in doing things* and *Feeling down, depressed or hopeless.* Both items are scored on a 4-point Likert-type scale ranging from 0 (*not at all*) to 3 (*nearly every day*). Scores range from 0 to 6, with a score of 3 or greater as the cutoff for a positive depression screen. The sensitivity and specificity of the PHQ–2 for identifying major depressive disorder are .83 and .90, respectively. The internal consistency of the PHQ–2 in this study was $r = .87$, and the total score from the PHQ–2 was used as a measure of severity.

Posttraumatic stress disorder

The Primary Care Post Traumatic Stress Disorder Questionnaire (PC–PTSD; Prins et al., 2004), is a 4-item (yes–no) screener for PTSD. Each item is a combination of PTSD symptoms grounded in *DSM–IV* criteria. Scores range from 0 to 4 with a score of 3 or greater as the cutoff for a positive PTSD screen. The time frame used for rating each item is "in the past month." Internal consistency of the PC–PTSD for this study was $\alpha = .78$, and total scores were used for symptom severity.

Health care utilization

These measures were described in Study 1.

Data analysis

We calculated the means, standard deviations, and internal consistency reliability for the EMI scales. Means and standard deviations for depression severity, BPD symptoms, child maltreatment severity, PTSD symptoms, number of ER visits, and number of overnight hospitalizations were also calculated. Self-report data and early memory narratives were complete for all 72 protocols

included in this study. We calculated Pearson correlations between the EMI scale scores and other variables. We used mean scores for the EMI scales for all analyses. We again used the Westen and Rosenthal (2003) method for testing the construct validity of the three EMI factors. We ran $r_{contrast\text{-}CV}$ effect size correlations comparing the hypothesized and actual magnitudes of the three EMI factors with all outcome measures—psychopathology, severity of trauma symptoms, and health care utilization.

Results

No significant differences were found between men and women on any of the study variables in the suburban sample using *t* tests. Thus, the data from men and women were aggregated. All of the variables were normally distributed (Curran et al., 1996). The mean scores of the women and men on each of the measures are shown in Table 1. As a group, the sample had mild levels of depression and a moderate level of childhood sexual and physical abuse. The mean scores were not elevated for BPD or PTSD symptoms.

Reliability and validity of the EMI

EMI item means and corrected item-to-total correlations are reported in Table 2.

Internal consistencies (Cronbach's α) of the scales were Mental Health, .90; Narrative Coherence, .83; and Traumatic Experiences, .83. Interrater reliabilities (ICC; one-way random effect) with Spearman–Brown correction for double coding were Mental Health, .87; Narrative Coherence, .89; and Traumatic Experiences, .79. Interrater reliability reached the excellent range (ICC > .75) for all three EMI scales. Intercorrelations of the EMI were Mental Health and Narrative Coherence, $r = .38$, $p < .001$; Mental Health and Traumatic Experiences, $r = -.60$, $p < .001$; and Narrative Coherence and Traumatic Experiences, $r = -.16$, $p = .12$. As hypothesized, the Mental Health and Traumatic Experiences scores were significantly correlated with both measures of psychopathology in the expected direction with medium effect sizes. The Mental Health score was negatively correlated with both measures of childhood abuse with small and medium effect sizes. The Mental Health score was not significantly correlated with a measure of PTSD symptoms or either measure of health care utilization. The Traumatic Experiences score was positively and significantly correlated with childhood physical abuse approaching a large effect size, PTSD symptoms with a medium effect size, and both measures of health care utilization with medium effect sizes. The Narrative Coherence scale negatively and significantly correlated with depression severity with a medium effect size and childhood sexual abuse with a small effect size. The correlations are shown in Table 3.

Validity coefficients from the $r_{contrast\text{-}CV}$ analyses for the suburban sample are also reported in Table 3. The magnitude of the $r_{contrast\text{-}CV}$ coefficient in the suburban sample indicates that the construct validity of the Mental Health score was good, the $r_{contrast\text{-}CV}$ coefficient for Traumatic Experiences score was fair, and the $r_{contrast\text{-}CV}$ coefficient for the Narrative Coherence score was fair.

Fisher *r*-to-*z* comparisons were calculated for correlations between Mental Health and Traumatic Experiences and measures of psychopathology, severity of trauma symptoms, and ER visits. Only the comparisons between Mental Health and hospital overnights ($r = -.07$) and Traumatic Experiences and hospital overnights ($r = .39$) differed statistically, $p = .04$.

Discussion

In both samples, the items of all three EMI scales demonstrated good item-to-total correlations, good levels of internal consistency, and excellent levels of interrater reliability. Despite using different numbers of early memories in each sample—10 in the urban sample and 5 in the suburban sample—the findings from both primary care sites were strikingly similar for the EMI Mental Health score. The utility of the EMI Mental Health scale was supported through robustly similar findings in two samples with different socioeconomic levels. The significant negative correlations between the Mental Health scale and measures of psychopathology (depression and BPD symptoms) support the convergent validity of the scale. This was further supported through $r_{contract\text{-}CV}$ analyses for the entire pattern of correlations between the EMI Mental Health score, psychopathology, severity of trauma symptoms, and health care utilization. As expected, a similar pattern of positive correlations was obtained between the Traumatic Experiences score and measures of psychopathology. Also as expected, the Mental Health score in both samples significantly and negatively correlated with measures of trauma. Mental Health significantly and negatively correlated with both measures of childhood abuse and with PTSD symptoms in the urban sample and showed a trend toward significance ($p = .09$) with PTSD symptoms in the suburban sample. The Traumatic Experiences score significantly and positively correlated with trauma measures. Although the correlations between the Traumatic Experiences score and measures of severity of trauma symptoms tended to be larger than the correlations between Traumatic Experiences score and measures of psychopathology, they were not significantly different. The $r_{contract\text{-}CV}$ analyses for the Traumatic Experiences score suggest that the validity of the Mental Health score is stronger than that of the Traumatic Experiences score, especially within the suburban sample. Further research is needed to understand the differential relationships between traumatic early memories and measures of trauma and psychopathology. The findings are surprisingly robust when considering that Study 1 and Study 2 used a different number and type of early memories, short and long versions of the PHQ depression scale, and different PTSD measures. Perhaps most important, the samples differed in gender and ethnic composition.

The Narrative Coherence score failed to show consistent associations with measures of psychopathology and was virtually unrelated to severity of trauma symptoms (only one significant correlation) and measures of health care utilization. Unlike previous studies showing changes in narrative coherence from the beginning to the end of PTSD treatment (Foa et al., 1995;

Van Minnen et al., 2002), ratings were based on portions of psychotherapy transcripts that might provide a greater opportunity to observe interruptions within the flow of narratives involving traumatic experiences. Likewise, narrative coherence ratings from the Adult Attachment Interview (AAI; George, Kaplan, & Main, 1985) are also based on longer narratives. Thus, the validity coefficients of the EMI Narrative Coherence score might be a function of the abbreviated length of early memory narratives.

In several EMI studies in the literature, Mental Health was negatively associated with health-related measures—rate-pressure product (a combination of heart rate and systolic blood pressure; Nehrig, 2014; Shedler et al., 2003), number of visits to a university health clinic, and verified physical illness (Cousineau & Shedler, 2006). In Cousineau and Shedler (2006), the Mental Health score was not associated with self-reported health care utilization. Although we did not have access to physiologic or verifiable illness data in either sample, we did have utilization data (ER use and overnight hospitalizations) from the MHP–H. The findings were different in the two primary care samples. The Mental Health score was significantly and negatively correlated with overnight hospitalizations in the urban but not the suburban sample. The Traumatic Experiences score was significantly correlated with overnight hospitalizations in both the urban and suburban samples. However, the Traumatic Experiences score was significantly correlated with ER use in only the suburban sample. Because overnight hospitalizations are driven more by verifiable physical pathology than ER use (i.e., a person will not usually be admitted to the hospital without verifiable signs of physical illness), it appears that the EMI Traumatic Experiences scale is more robustly associated with physical health status in primary care samples than the Mental Health scale. The relationship between early trauma and adult physical health problems has been well documented (e.g., Felitti & Anda, 2010).

There are several limitations in these studies. First, there was an overrepresentation of women in both studies, which limits the generalizability of the findings to men using primary care. Future research should improve on the ratio of women to men, despite the fact that women make up approximately two thirds of all primary care patients (U.S. Department of Health and Human Services, Health Resources and Services Administration, 2013). Second, both studies lacked measures for the assessment of discriminant validity. Finally, the studies used different numbers and types of early memories, different forms of the PHQ depression scales, and different measures of PTSD symptoms. Despite these limitations, the study provides support for the use of the Mental Health scale and partial support for the Traumatic Experiences and Narrative Coherence scales of the EMI for use with primary care patients.

Funding

Study 1 (urban primary care study) was partially supported by federal matching funds through the Michigan Department of Community Health and the Wayne State University Department of Family Medicine (MDCH 224–23M2Q) awarded to John H. Porcerelli. Study 2 (suburban primary care study) was supported by Family Medicine Residency faculty research funds.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bernstein, D., & Fink, L. (1998). *Manual for the Childhood Trauma Questionnaire*. San Antonio, TX: Psychological Corporation.

Bernstein, D., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., … Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse and Neglect*, *27*, 169–190.

Bruhn, A. R. (1985). Using early memories as a projective technique—The cognitive perceptual method. *Journal of Personality Assessment*, *49*, 587–597.

Bruhn, A. R. (1990). Cognitive-perceptual theory and the projective use of autobiographical memory. *Journal of Personality Assessment*, *55*, 95–114.

Carruthers, M. J. (1966). *The book of memory: A study of memory in medieval culture*. Cambridge, UK: Cambridge University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cousineau, T. M., & Shedler, J. (2006). Predicting physical health: Implicit mental health measures versus self-report scales. *The Journal of Nervous and Mental Disease*, *194*, 427–432.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*, 16–29.

Davison, S., Leese, M., & Taylor, P. J. (2001). Examination of the screening properties of the Personality Diagnostic Questionnaire 4+ (PDQ–4+) in a prison population. *Journal of Personality Disorders*, *15*, 180–194.

Felitti, V. J., & Anda, R. F. (2010). The relationship of adverse childhood experiences to adult medical disease, psychiatric disorders, and sexual behavior: Implications for healthcare. In R. A. Lanius, E. Vermetten, & C. Pain (Eds.), *The hidden epidemic: The impact of early life trauma on health and disease* (pp. 77–87). New York, NY: Cambridge University Press.

Foa, E. B., Molnar, C., & Cashman, L. (1995). Change in rape narratives during exposure therapy for posttraumatic stress disorder. *Journal of Traumatic Stress*, *8*, 675–690.

Fowler, J. C., Hilsenroth, M. J., & Handler, L. (1995). Early memories: An exploration of theoretically derived queries and their clinical utility. *Bulletin of the Menninger Clinic*, *59*, 79–98.

Freud, S. (1961). Screen memories. In J. Strachey (Ed. and Trans.), *The standard edition of the complete works of Sigmund Freud* (Vol. 3, pp. 303–322). London, UK: Hogarth Press. (Original work published 1899)

George, C., Kaplan, N., & Main, M. (1985). *Adult Attachment Interview*. Unpublished manuscript, University of California, Berkeley.

Gross, R., Olfson, M., Gameroff, M., Shae, S., Feder, A., Fuentes, M., … Weissman, M. M. (2002). Borderline personality disorder in primary care. *Archives of Internal Medicine*, *162*, 53–60.

Hyler, S. E. (1994). *Personality Diagnostic Questionnaire, DSM–IV version (PDQ–4+)*. New York, NY: New York State Psychiatric Institute.

Karliner, R., Westrich, E. K, Shedler, J., & Mayman, M. (1996). Bridging the gap between psychodynamic and scientific psychology: The Adelphi Early Memory Index. In J. M. Masling & R. F. Bornstein (Eds.), *Empirical studies of psychoanalytic theories: Psychoanalytic perspectives on developmental psychology* (pp. 43–67). Washington, DC: American Psychological Association.

Karoly, P., Ruehlman, L. S., & Lanyon, R. I. (2005). The assessment of adult health care orientations: Development and preliminary validation of the Multidimensional Health Profile–Health Functioning Index (MHP–H) in a national sample. *Journal of Clinical Psychology in Medical Settings*, *12*, 79–91.

Katon, W., & Schulberg, H. (1992). Epidemiology of depression in primary care. *General Hospital Psychiatry, 14*, 237–247.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). PHQ–9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606–616.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The Patient Health Questionnaire–2: Validity of a two item depression screener. *Medical Care*, *41*, 1284–1294.

Langs, R. (1965a). Earliest memories and personality: A predictive study. *Archives of General Psychiatry*, *18*, 379–390.

Langs, R. (1965b). First memories and characterologic diagnosis. *Journal of Nervous and Mental Disease*, *141*, 318–320.

Langs, R. J., Rothenberg, M. B., Fishman, J. R., & Reiser, M. F. (1960). A method for clinical and theoretical study of early memory. *Archives of General Psychiatry*, *3*, 523–534.

Lin, O. S. P. (2006). *The early spiritual memory test: A psychometric evaluation of a new measure* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3232480)

Mayman, M. (1963). Psychoanalytic study of the self-organization with psychological tests. In B. T. Wigdor (Ed.), *Recent advances in the study of behavior change: Proceedings of the Academic Academy on Clinical Psychology* (pp. 97–117). Montreal, Canada: McGill University Press.

Mayman, M. (1968). Early memories and character structure. *Journal of Projective Techniques and Personality Assessment*, *32*, 303–316.

Mayman, M., & Faris, M. (1960). Early memories as expressions of relationship paradigms. *American Journal of Orthopsychiatry*, *30*, 507–520.

Miles, C. (1895). A study of individual psychology. *American Journal of Psychology*, *6*, 534–558.

Nehrig, N. (2014). *Self-reported mental health of narcissists: Illusion or reality?* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3579848)

Okada, M., & Oltmanns, T. F. (2009). Comparison of three self-report measures of personality pathology. *Journal of Psychopathology and Behavioral Assessment, 31*, 358–367.

Porcerelli, J. H., Cogan, R., Markova, T., Murdoch, W., & Porcerelli, M. (2010). Abuse, outpatient charges and utilization, and psychiatric symptoms among urban women on Medicaid. *Journal of the American Board of Family Medicine*, *23*, 363–370.

Porcerelli, J. H., & Huprich, S. (2007). Approach to personality disorders in primary care. *Hospital Physician*, *43*, 3–11.

Prins, A., Ouimette, P., Kimerling, R., Cameron, D.S. Hugelshofer, J., & Shaw-Hegwer, J. (2004). The primary care PTSD screen (PC–PTSD): development and operating characteristics. *Primary Care Psychiatry*, *9*, 9–14.

Ruehlman, L. S., Lanyon, R. I., & Karoly, P. (1998). *Multidimensional Health Profile: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Ruggiero, K. J., Del Ben, K., Scotti, J. R., & Rabalais, A. E. (2003). Psychometric properties of the PTSD Checklist–Civilian version. *Journal of Traumatic Stress*, *16*, 495–502.

Schrecker, P. (1973). Individual psychological significance of first childhood recollections. *Journal of Individual Psychology*, *29*, 146–156. (Original work published 1913)

Shedler, J., Karliner, R., & Katz, E. (2003). Cloning the clinician: A method for assessing illusory mental health. *Journal of Clinical Psychology*, *59*, 635–650.

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

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlation: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428.

Spitzer R. L., Kroenke, K., & Williams, J. B. W. (1999). Validation and utility of a self-report version of PRIME-MD. *Journal of the American Medical Association*, *282*, 1737–1744.

Spitzer, R. L., Williams, J. B. W., Kroenke, K., Hornyak, R., & McMurray, J. (2000). Validity and utility of the PRIME-MD Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD Patient Health Questionnaire Obstetrics-Gynecologic study. *American Journal of Obstetrics and Gynecology*, *183*, 759–769.

Taylor, S. E., Lerner, J. S., Sherman, D. K., Sage, R. M., & McDowell, N. K. (2003). Portrait of the self-enhancer: Well-adjusted and well liked or maladjusted and friendless? *Journal of Personality and Social Psychology*, *84*, 165–176.

U.S. Department of Health and Human Services, Health Resources and Services Administration. (2013). *Women's health USA 2013*. Rockville, MD: Author. Retrieved from http://mchb.hrsa.gov/whusa13/dl/pdf/whusa13.pdf

Van Minnen, A., Wessel, I., Dijkstra, T., & Roelofs, K. (2002). Changes in PTSD patients' narratives during prolonged exposure therapy: A replication and extension. *Journal of Traumatic Stress*, *15*, 255–258.

Weathers, F. W., Litz, B. T., Herman, J. A., Huska, J. A., & Keane, T. M. (1993, October). *The PTSD checklist (PCL): Reliability, validity, and diagnostic utility*. Paper presented at the meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology*, *84*, 608–618.